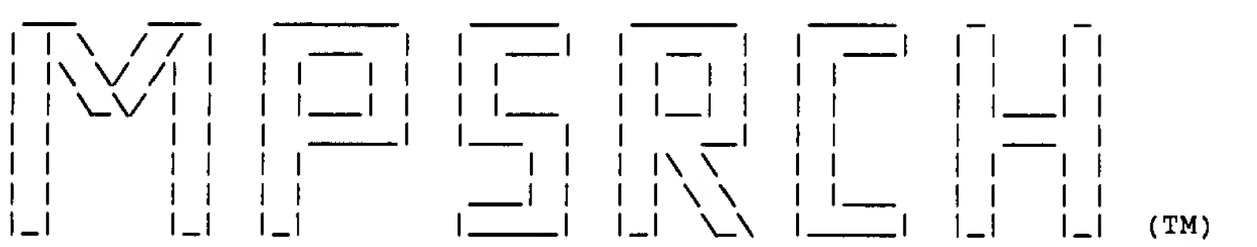


MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Oct 22 15:56:22 1998; MasPar time 8.60 Seconds
801.905 Million cell updates/sec
Tabular output not generated.

Title: >US-08-848-439-3
Description: (1-275) from US08848439.pep
Perfect Score: 2073
Sequence: 1 SARGLFLFGQPDFSYKRSNC.....WQKGQREFKRISRSIRKLQC 275

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 46.355; Variance 72.254; scale 0.642

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	332	16.0	581	1	FRIZ_DROME	FRIZZLED PROTEIN PRECU	5.63e-52
2	148	7.1	581	1	NET2_CHICK	NETRIN-2 PRECURSOR (FR	1.06e-10
3	124	6.0	606	1	NET1_CHICK	NETRIN-1 PRECURSOR.	3.47e-06
4	109	5.3	465	1	YQI8_CAEEL	HYPOTHETICAL 53.8 KD P	1.39e-03
5	101	4.9	1741	1	RPC1_GIALA	DNA-DIRECTED RNA POLYM	2.78e-02
6	100	4.8	500	1	CPDI_RAT	CYTOCHROME P450 IID18	3.99e-02
7	100	4.8	500	1	CPD4_RAT	CYTOCHROME P450 IID4 (	3.99e-02
8	99	4.8	724	1	KS61_MOUSE	RIBOSOMAL PROTEIN S6 K	5.72e-02
9	98	4.7	1032	1	VG07_BPT4	BASEPLATE STRUCTURAL P	8.18e-02
10	96	4.6	520	1	ION3_CARAU	INTERMEDIATE FILAMENT	1.66e-01
11	94	4.5	531	1	UD13_RAT	UDP-GLUCURONOSYLTRANSF	3.32e-01
12	92	4.4	1562	1	YM81_YEAST	HYPOTHETICAL 180.2 KD	6.58e-01
13	90	4.3	216	1	GYRB_ACIS3	DNA GYRASE SUBUNIT B (	1.29e+00
14	89	4.3	262	1	VIRF_SHIDY	VIRULENCE REGULON TRAN	1.79e+00
15	90	4.3	263	1	KSGA_MYCPN	DIMETHYLADENOSINE TRAN	1.29e+00
16	90	4.3	561	1	VATA_MAIZE	VACUOLAR ATP SYNTHASE	1.29e+00
17	89	4.3	752	1	KS6A_CHICK	RIBOSOMAL PROTEIN S6 K	1.79e+00
18	88	4.2	127	1	YBU9_YEAST	HYPOTHETICAL 14.8 KD P	2.48e+00
19	88	4.2	182	1	YLP2_CAEEL	HYPOTHETICAL 20.1 KD P	2.48e+00
20	88	4.2	443	1	AM2A_ORYSA	ALPHA-AMYLASE ISOZYME	2.48e+00
21	88	4.2	445	1	AMC2_ORYSA	ALPHA-AMYLASE ISOZYME	2.48e+00
22	87	4.2	446	1	YHE2_PSEAE	HYPOTHETICAL 50.0 KD P	3.43e+00
23	87	4.2	510	1	NOA1_HUMAN	ONCONEURAL VENTRAL ANT	3.43e+00
24	87	4.2	580	1	VATA_HORVU	VACUOLAR ATP SYNTHASE	3.43e+00
25	88	4.2	633	1	KS62_MOUSE	RIBOSOMAL PROTEIN S6 K	2.48e+00
26	88	4.2	740	1	KS62_HUMAN	RIBOSOMAL PROTEIN S6 K	2.48e+00
27	88	4.2	848	1	NFM_MOUSE	NEUROFILAMENT TRIPLET	2.48e+00
28	88	4.2	1102	1	YK64_CAEEL	HYPOTHETICAL 124.8 KD	2.48e+00
29	87	4.2	1201	1	COPA_YEAST	COATOMER ALPHA SUBUNIT	3.43e+00
30	87	4.2	1752	1	DESP_HUMAN	DESMOPLAKIN I AND II (	3.43e+00
31	87	4.2	2749	1	IP3R_MOUSE	INOSITOL 1,4,5-TRISPHO	3.43e+00
32	87	4.2	2749	1	IP3R_RAT	INOSITOL 1,4,5-TRISPHO	3.43e+00
33	85	4.1	216	1	GYRB_ACIS7	DNA GYRASE SUBUNIT B (	6.48e+00
34	86	4.1	327	1	VMP_CAMVD	MOVEMENT PROTEIN (CELL	4.72e+00
35	86	4.1	450	1	DCOR_CHICK	ORNITHINE DECARBOXYLAS	4.72e+00
36	86	4.1	461	1	DCOR_BOVIN	ORNITHINE DECARBOXYLAS	4.72e+00
37	85	4.1	495	1	E1BL_ADE02	E1B PROTEIN, LARGE T-A	6.48e+00
38	86	4.1	556	1	FTHS_STRMU	FORMATE--TETRAHYDROFOL	4.72e+00
39	86	4.1	620	1	Y870_METJA	HYPOTHETICAL PROTEIN M	4.72e+00
40	85	4.1	633	1	YHR4_YEAST	HYPOTHETICAL 71.2 KD P	6.48e+00
41	86	4.1	725	1	VACB_MYCGE	VACB PROTEIN HOMOLOG.	4.72e+00
42	85	4.1	798	1	VP16_YEAST	VACUOLAR PROTEIN SORTI	6.48e+00
43	85	4.1	1407	1	CYAA_DICDI	ADENYLATE CYCLASE, AGG	6.48e+00
44	85	4.1	1726	1	MSP1_PLAFP	MEROZOITE SURFACE PROT	6.48e+00
45	85	4.1	2688	1	ZEP1_MOUSE	ZINC FINGER PROTEIN 40	6.48e+00

ALIGNMENTS

RESULT 1
ID FRIZ_DROME STANDARD; PRT; 581 AA.
AC P18537;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE FRIZZLED PROTEIN PRECURSOR.
GN FZ.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN [1]
RP SEQUENCE FROM N.A. (CLASS I).
RX MEDLINE; 89159415.
RA VINSON C.R., CONOVER S., ADLER P.N.;
RL NATURE 338:263-264(1989).
RN [2]
RP SEQUENCE FROM N.A. (CLASSES I AND II).
RX MEDLINE; 91060073.
RA ADLER P.N., VINSON C., PARK W.J., CONOVER S., KLEIN L.;
RL GENETICS 126:401-416(1990).
CC -!- FUNCTION: REQUIRED TO COORDINATE THE CYTOSKELETONS OF EPIDERMAL
CC CELLS TO PRODUCE A PARALLEL ARRAY OF CUTICULAR HAIRS AND BRISTLES.
CC FZ IS REQUIRED FOR THE TRANSMISSION OF POLARITY INFORMATION, AND
CC TRANSDUCTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: CLASS I AND CLASS II PROTEINS DIFFER AT THE
CC C-TERMINAL FROM AA 406 DUE TO USE OF ALTERNATIVE 3' EXONS.
DR EMBL; X54648; G804979; -.
DR EMBL; X54649; G804979; JOINED.
DR EMBL; X54650; G804979; JOINED.
DR EMBL; X54651; G804979; JOINED.
DR EMBL; X54648; G804980; -.
DR EMBL; X54649; G804980; JOINED.
DR EMBL; X54650; G804980; JOINED.
DR EMBL; X54652; G804980; JOINED.
DR EMBL; X54646; G7981; -.
DR EMBL; X54647; G7983; ALT_SEQ.
DR PIR; S03540; S03540.
DR PIR; S15708; S15708.
DR FLYBASE; FBgn0001085; fz.
KW TRANSMEMBRANE; SIGNAL; ALTERNATIVE SPLICING; DEVELOPMENTAL PROTEIN.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 581 FRIZZLED PROTEIN.
FT TRANSMEM 248 270 POTENTIAL.
FT TRANSMEM 281 303 POTENTIAL.
FT TRANSMEM 336 368 POTENTIAL.
FT TRANSMEM 381 401 POTENTIAL.

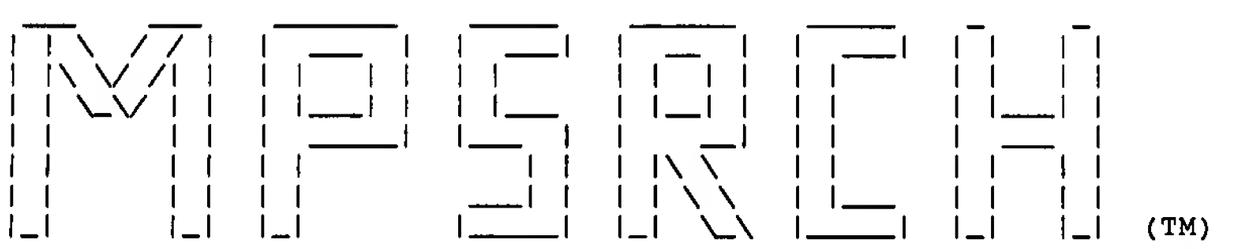
```
FT   TRANSMEM     422     439       POTENTIAL.
FT   TRANSMEM     472     492       POTENTIAL.
FT   TRANSMEM     530     553       POTENTIAL.
FT   VARSPLIC     406     415       DILSGVCFVG -> MYLWQFHTIN (IN CLASS II).
FT   VARSPLIC     416     581       MISSING (IN CLASS II).
SQ   SEQUENCE   581 AA;  64847 MW;  279EBE85 CRC32;

  Query Match           16.0%;  Score 332;  DB 1;  Length 581;
  Best Local Similarity  37.5%;  Pred. No. 5.63e-52;
  Matches    45;  Conservative   22;  Mismatches  47;  Indels   6;  Gaps   6;

Db      50 HNRCEPI-T-ISICKNIPYNMTIMPNLIGHTKQEEAGLEVHQFAPLVKIGCSDDLQLFLC 107
           :: | || : : :|  | |    :|||:||    |   :     |||:  |  |   |||
Qy      17 RSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 76

Db     108 SLYVPVC-TILERPIPPCRSLCESARV-CEKLMKTYNFNWPENLECSKFPVHGGEDLCVA 165
           ||: |||   |:  | ||:|||   :  |  :| :: | ||: ||| :|| :  :|||::
Qy      77 SLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QD-NDLCIP 134


RESULT    2
ID   NET2_CHICK     STANDARD;      PRT;   581 AA.
AC   Q90923;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   NETRIN-2 PRECURSOR (FRAGMENT).
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WHITE LEGHORN; TISSUE=EMBRYONIC BRAIN;
RX   MEDLINE; 94340732.
RA   SERAFINI T., KENNEDY T.E., GALKO M.J., MIRZAYAN C., JESSELL T.M.,
RA   TESSIER-LAVIGNE M.;
RL   CELL 78:409-424(1994).
CC   -!- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND
CC       PERIPHERAL MOTOR AXONS.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC   -!- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS ONE C345C DOMAIN.
CC   -!- SIMILARITY: TO CAENORHABDITIS ELEGANS UNC-6 PROTEIN.
DR   EMBL; L34550; G529421; -.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 3.
KW   GLYCOPROTEIN; EXTRACELLULAR MATRIX; SIGNAL; LAMININ EGF-LIKE DOMAIN;
KW   REPEAT.
FT   NON_TER       1       1
FT   SIGNAL       <1      15       POTENTIAL.
FT   CHAIN         16     581       NETRIN-2.
FT   DOMAIN        16     261       LAMININ N-TERMINAL (DOMAIN VI).
FT   DOMAIN       262     430       3 X LAMININ EGF-LIKE REPEATS (LAMININ
FT                                  DOMAIN V-LIKE).
FT   DOMAIN       262     317       LAMININ EGF-LIKE 1.
FT   DOMAIN       318     380       LAMININ EGF-LIKE 2.
FT   DOMAIN       381     430       LAMININ EGF-LIKE 3.
FT   DOMAIN       431     581       C345C (DOMAIN C).
FT   SITE         507     509       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID      91     124       BY SIMILARITY.
FT   DISULFID     262     271       BY SIMILARITY.
FT   DISULFID     264     281       BY SIMILARITY.
FT   DISULFID     283     292       BY SIMILARITY.
FT   DISULFID     295     315       BY SIMILARITY.
FT   DISULFID     318     327       BY SIMILARITY.
FT   DISULFID     320     345       BY SIMILARITY.
FT   DISULFID     348     357       BY SIMILARITY.
FT   DISULFID     360     378       BY SIMILARITY.
FT   DISULFID     381     393       BY SIMILARITY.
FT   DISULFID     383     400       BY SIMILARITY.
FT   DISULFID     402     411       BY SIMILARITY.
FT   DISULFID     414     428       BY SIMILARITY.
FT   DISULFID     453     521       BY SIMILARITY.
FT   DISULFID     468     578       BY SIMILARITY.
FT   CARBOHYD      88      88       POTENTIAL.
FT   CARBOHYD     103     103       POTENTIAL.
FT   CARBOHYD     394     394       POTENTIAL.
FT   CARBOHYD     540     540       POTENTIAL.
SQ   SEQUENCE   581 AA;  65106 MW;  5B6D2272 CRC32;

  Query Match            7.1%;  Score 148;  DB 1;  Length 581;
  Best Local Similarity  27.0%;  Pred. No. 1.06e-10;
  Matches    30;  Conservative   28;  Mismatches  46;  Indels   7;  Gaps   6;

Db     431 IPAINPTSLVTSTE-APADCDSYCKPAKGNYKINMKKYCKKDYVVQVNILEMETVANWAK 489
           ||  ::  |: :|| ||  |:: ||  : : :  |   ||:|: : ::: |:  :   :|
Qy     133 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTK 191

Db     490 FTINILS--VYKCRDERVKR-GDNFLWIHLKDLSCKCPKI-QISKKYLVMG 536
           : ::  |  :||      :    : ||:    | | |  : :|:  |||||
Qy     192 IILETKSKTIYKLNGVSERDLKKSVLWLK-DSLQCTCEEMNDINAPYLVMG 241


RESULT    3
ID   NET1_CHICK     STANDARD;      PRT;   606 AA.
AC   Q90922;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   NETRIN-1 PRECURSOR.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WHITE LEGHORN; TISSUE=EMBRYONIC BRAIN;
RX   MEDLINE; 94340732.
RA   SERAFINI T., KENNEDY T.E., GALKO M.J., MIRZAYAN C., JESSELL T.M.,
RA   TESSIER-LAVIGNE M.;
RL   CELL 78:409-424(1994).
CC   -!- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND
CC       PERIPHERAL MOTOR AXONS.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC   -!- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS ONE C345C DOMAIN.
CC   -!- SIMILARITY: TO CAENORHABDITIS ELEGANS UNC-6 PROTEIN.
DR   EMBL; L34549; G529419; -.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 3.
KW   GLYCOPROTEIN; EXTRACELLULAR MATRIX; SIGNAL; LAMININ EGF-LIKE DOMAIN;
KW   REPEAT.
FT   SIGNAL        1      25       POTENTIAL.
FT   CHAIN        26     606       NETRIN-1.
FT   DOMAIN       26     286       LAMININ N-TERMINAL (DOMAIN VI).
FT   DOMAIN      287     455       3 X LAMININ EGF-LIKE REPEATS (LAMININ
FT                                  DOMAIN V-LIKE).
FT   DOMAIN      287     342       LAMININ EGF-LIKE 1.
FT   DOMAIN      343     405       LAMININ EGF-LIKE 2.
FT   DOMAIN      406     455       LAMININ EGF-LIKE 3.
FT   DOMAIN      456     606       C345C (DOMAIN C).
FT   SITE        532     534       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    121     154       BY SIMILARITY.
FT   DISULFID    287     296       BY SIMILARITY.
FT   DISULFID    289     306       BY SIMILARITY.
FT   DISULFID    308     317       BY SIMILARITY.
FT   DISULFID    320     340       BY SIMILARITY.
FT   DISULFID    343     352       BY SIMILARITY.
FT   DISULFID    345     370       BY SIMILARITY.
FT   DISULFID    373     382       BY SIMILARITY.
FT   DISULFID    385     403       BY SIMILARITY.
FT   DISULFID    406     418       BY SIMILARITY.
FT   DISULFID    408     425       BY SIMILARITY.
FT   DISULFID    427     436       BY SIMILARITY.
FT   DISULFID    439     453       BY SIMILARITY.
```

```
FT   DISULFID    478    546       BY SIMILARITY.
FT   DISULFID    493    603       BY SIMILARITY.
FT   CARBOHYD     97     97       POTENTIAL.
FT   CARBOHYD    118    118       POTENTIAL.
FT   CARBOHYD    133    133       POTENTIAL.
FT   CARBOHYD    419    419       POTENTIAL.
SQ   SEQUENCE   606 AA;  68126 MW;  C07CCA06 CRC32;

  Query Match             6.0%;  Score 124;  DB 1;  Length 606;
  Best Local Similarity  26.1%;  Pred. No. 3.47e-06;
  Matches    31;  Conservative   34;  Mismatches  41;  Indels  13;  Gaps  11;

Db    456 IPAAPPPTAASSTEEPAD-CDSYCKASKGKLKINMKKYCKKDYAVQIHILKAEKNADW-W 513
          || |::    ::|||::  |:: ||  :   :  |   ||:|:|: | : |     ::
Qy    133 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDT 190

Db    514 KFTVNIIS--VYK-QG-SNR-LRRGDQTLWVHAKDIACKCPKVKPMKK-YLLLGSTEDS 566
          |: ::  |  :||  | |:| |:::   ||:    : | |  :: ::  ||::|  : :
Qy    191 KIILETKSKTIYKLNGVSERDLKKS--VLWLKD-SLQCTCEEMNDINAPYLVMGQKQGG 246


RESULT    4
ID   YQI8_CAEEL     STANDARD;      PRT;   465 AA.
AC   Q09280;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 53.8 KD PROTEIN C45G9.8 IN CHROMOSOME III.
GN   C45G9.8.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   BENTLEY D., WATERSTON R.;
RL   SUBMITTED (FEB-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U21323; G687873; -.
DR   WORMPEP; C45G9.8; CE01859.
KW   HYPOTHETICAL PROTEIN.
FT   DOMAIN      139    145       POLY-THR.
SQ   SEQUENCE   465 AA;  53756 MW;  37C8121B CRC32;

  Query Match             5.3%;  Score 109;  DB 1;  Length 465;
  Best Local Similarity  33.3%;  Pred. No. 1.39e-03;
  Matches    25;  Conservative   19;  Mismatches  25;  Indels   6;  Gaps   6;

Db    202 QLRPRIEEDQNI-EV-TSRDIDEAD-MERICKK-FAKEKINPITMAEPIDEKTKILLDRV 257
          :| |  ||  :: |   ::: |: | || :||: ||  ||:   ::  |:  |||:|:
Qy    140 HLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFAL-KIKVKEITY-INRDTKIILETK 197

Db    258 TKKPYPLKYNNEKGL 272
          :|  | |:  :|: |
Qy    198 SKTIYKLNGVSERDL 212


RESULT    5
ID   RPC1_GIALA     STANDARD;      PRT;  1741 AA.
AC   P25202;
DT   01-MAY-1992 (REL. 22, CREATED)
DT   01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   DNA-DIRECTED RNA POLYMERASE III LARGEST SUBUNIT (EC 2.7.7.6).
GN   RPOA3.
OS   GIARDIA LAMBLIA (GIARDIA INTESTINALIS).
OC   EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; MASTIGOPHORA; DIPLOMONADIDA;
OC   HEXAMITIDAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SSP. P1;
RX   MEDLINE; 92195823.
RA   LANZENDOERFER M., PALM P., GRAMPP B., PEATTIE D.A., ZILLIG W.;
RL   NUCLEIC ACIDS RES. 20:1145-1145(1992).
CC   -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC       OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC       SUBSTRATES.
CC   -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
CC       RNA(N).
CC   -!- SUBUNIT: RNA POLYMERASE III CONSISTS OF ABOUT 15 DIFFERENT
CC       SUBUNITS. THIS SUBUNIT IS THE LARGEST COMPONENT OF RNA POLYMERASE
CC       III.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE FOUND IN
CC       EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR,
CC       POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE III FOR 5S
CC       AND TRNA GENES.
DR   EMBL; X60325; G9364; -.
KW   DNA-DIRECTED RNA POLYMERASE; TRANSCRIPTION; ZINC; ZINC-FINGER;
KW   NUCLEAR PROTEIN.
FT   ZN_FING      79     92       C2H2-TYPE (POTENTIAL).
SQ   SEQUENCE   1741 AA;  193928 MW;  DCDCD942 CRC32;

  Query Match             4.9%;  Score 101;  DB 1;  Length 1741;
  Best Local Similarity  27.4%;  Pred. No. 2.78e-02;
  Matches    20;  Conservative   25;  Mismatches  23;  Indels   5;  Gaps   5;

Db    168 TICKKSQGFRV-LREMHYEERATKILADVIKHQKDIEKIQELDQR-LRNHVVDPIQALHI 225
          |:||:  :::: ::|: | :| |||: :  |  | | |:  : :| |:: |:   ::|:
Qy    168 TLCKNDFALKIKVKEITYINRDTKIILET-KS-KTIYKLNGVSERDLKKSVLWLKDSLQC 225

Db    226 L-QKVPECDYPYL 237
            : : : : |||
Qy    226 TCEEMNDINAPYL 238


RESULT    6
ID   CPDI_RAT       STANDARD;      PRT;   500 AA.
AC   Q64680;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   CYTOCHROME P450 IID18 (EC 1.14.14.1) (P450 2D-29/2D-35).
GN   CYP2D18 OR CYP2D-18.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95251650.
RA   KAWASHIMA H., STROBEL H.W.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 209:535-540(1995).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC       OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC       ACIDS, AND XENOBIOTICS.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC   -!- TISSUE SPECIFICITY: BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   EMBL; U48220; G1200518; -.
DR   EMBL; U48219; G1200516; -.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   OXIDOREDUCTASE; MONOOXYGENASE; ELECTRON TRANSPORT; MEMBRANE; HEME;
KW   MICROSOME.
FT   BINDING     446    446       HEME (BY SIMILARITY).
SQ   SEQUENCE   500 AA;  56683 MW;  F01949B5 CRC32;

  Query Match             4.8%;  Score 100;  DB 1;  Length 500;
  Best Local Similarity  26.7%;  Pred. No. 3.99e-02;
  Matches    16;  Conservative   22;  Mismatches  18;  Indels   4;  Gaps   3;

Db    173 FPFS-PNTLLDKAVCNVIASLLFACR-FEYNDPRFIRLL--DLLKDTLEEESGFLPMLLN 228
          | |: |:    :: |: |:: |  |: :|| : |:  ||  : :|: ||: :: :|::::
Qy      6 FLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMK 65
```

```
RESULT    7
ID   CPD4_RAT       STANDARD;      PRT;   500 AA.
AC   P13108;
DT   01-JAN-1990 (REL. 13, CREATED)
DT   01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   CYTOCHROME P450 IID4 (EC 1.14.14.1) (P450-DB4) (P450-CMF3)
DE   (DEBRISOQUINE 4-HYDROXYLASE).
GN   CYP2D14 OR CYP2D-4.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX   MEDLINE; 90189185.
RA   MATSUNAGA E., UMENO M., GONZALEZ F.J.;
RL   J. MOL. EVOL. 30:155-169(1990).
RN   [2]
RP   SEQUENCE OF 177-500 FROM N.A.
RX   MEDLINE; 89050091.
RA   ISHIDA N., TAWARAGI Y., INUZUKA C., SUGITA O., KUBOTA I.,
RA   NAKAZATO H., NOGUCHI T., SASSA S.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 156:681-688(1988).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC       OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC       ACIDS, AND XENOBIOTICS.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC   -!- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER
CC       TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,
CC       AND CARCINOGENS.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   EMBL; X52029; G57816; -.
DR   EMBL; M22331; G203830; -.
DR   PIR; D31579; D31579.
DR   PIR; S16873; S16873.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   OXIDOREDUCTASE; MONOOXYGENASE; ELECTRON TRANSPORT; MEMBRANE; HEME;
KW   MICROSOME.
FT   BINDING     446    446       HEME (BY SIMILARITY).
SQ   SEQUENCE   500 AA;  56697 MW;  429F82E7 CRC32;

  Query Match             4.8%;  Score 100;  DB 1;  Length 500;
  Best Local Similarity  26.7%;  Pred. No. 3.99e-02;
  Matches    16;  Conservative   22;  Mismatches  18;  Indels   4;  Gaps   3;

Db    173 FPFS-PNTLLDKAVCNVIASLLFACR-FEYNDPRFIRLL--DLLKDTLEEESGFLPMLLN 228
          | |: |:    :: |: |:: |  |: :|| : |:  ||  : :|: ||: :: :|::::
Qy      6 FLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMK 65


RESULT    8
ID   KS61_MOUSE     STANDARD;      PRT;   724 AA.
AC   P18653;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   RIBOSOMAL PROTEIN S6 KINASE II ALPHA 1 (EC 2.7.1.-) (S6KII-ALPHA)
DE   (P90-RSK 1).
GN   RPS6KA1 OR RSK1.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89384612.
RA   ALCORTA D.A., CREWS C.M., SWEET L.J., BANKSTON L., JONES S.W.,
RA   ERIKSON R.L.;
RL   MOL. CELL. BIOL. 9:3850-3859(1989).
CC   -!- FUNCTION: PHOSPHORYLATES A WIDE RANGE OF SUBSTRATES INCLUDING
CC       RIBOSOMAL PROTEIN S6.
CC   -!- TISSUE SPECIFICITY: INTESTINE, THYMUS, AND LUNG.
CC   -!- SIMILARITY: THE N-TERMINUS SHOWS HIGH SEQUENCE SIMILARITY TO THE
CC       REGIONS OF PROTEIN KINASE C, THE CATALYTIC SUBUNIT OF CAMP-
CC       DEPENDENT PROTEIN KINASE, AND CGMP-DEPENDENT PROTEIN KINASE. THE
CC       REMINDER OF S6 KINASE II SHOWS HIGH SEQUENCE SIMILARITY TO THE
CC       CATALYTIC SUBUNIT OF PHOSPHORYLASE B KINASE.
DR   EMBL; M28489; G556322; -.
DR   PIR; B32571; B32571.
DR   HSSP; P05132; 1CTP.
DR   MGD; MGI:104558; RPS6KA1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 2.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 2.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 2.
KW   TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW   REPEAT; MULTIGENE FAMILY.
FT   DOMAIN       62    310       PROTEIN KINASE 1.
FT   DOMAIN      407    664       PROTEIN KINASE 2.
FT   NP_BIND      68     76       ATP (BY SIMILARITY).
FT   BINDING      94     94       ATP (BY SIMILARITY).
FT   ACT_SITE    187    187       BY SIMILARITY.
FT   NP_BIND     413    421       ATP (BY SIMILARITY).
FT   BINDING     436    436       ATP (BY SIMILARITY).
FT   ACT_SITE    524    524       BY SIMILARITY.
SQ   SEQUENCE   724 AA;  81594 MW;  F961A261 CRC32;

  Query Match             4.8%;  Score 99;  DB 1;  Length 724;
  Best Local Similarity  35.2%;  Pred. No. 5.72e-02;
  Matches    19;  Conservative   13;  Mismatches  18;  Indels   4;  Gaps   4;

Db    489 KI-LRQKFFSEREASFVLHTISKTVEYLHSQGVVHRDLKPSNILYVDESGNPEC 541
          || :::  : :|:: ::| | |||: |    ||  |||| | :| : :|    |
Qy    177 KIKVKEITYINRDTKIILETKSKTI-Y-KLNGVSERDLKKS-VLWLKDSLQCTC 227


RESULT    9
ID   VG07_BPT4      STANDARD;      PRT;  1032 AA.
AC   P19061;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT   01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
DE   BASEPLATE STRUCTURAL PROTEIN GP7.
GN   7.
OS   BACTERIOPHAGE T4.
OC   VIRIDAE; DS-DNA NONENVELOPED VIRUSES; MYOVIRIDAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D;
RX   MEDLINE; 90384864.
RA   EFIMOV V.P., PRILIPOV A.G., MESYANZHINOV V.V.;
RL   NUCLEIC ACIDS RES. 18:5313-5313(1990).
CC   -!- FUNCTION: STRUCTURAL COMPONENT OF THE BASEPLATE.
DR   EMBL; X15907; G15323; -.
DR   PIR; JQ0657; G7BPT4.
KW   STRUCTURAL PROTEIN.
SQ   SEQUENCE  1032 AA;  119214 MW;  39F5B2D1 CRC32;

  Query Match             4.7%;  Score 98;  DB 1;  Length 1032;
  Best Local Similarity  39.0%;  Pred. No. 8.18e-02;
  Matches    16;  Conservative    8;  Mismatches  15;  Indels   2;  Gaps   2;

Db    368 MDSICDKVFALIIGEVETLNANPRTSKII-DSADKGIYVLN-407
          |:::| : ||| |  |  :    | :||| ::  | || ||
Qy    166 METLCKNDFALKI-KVKEITYINRDTKIILETKSKTIYKLN 205


RESULT   10
ID   ION3_CARAU     STANDARD;      PRT;   520 AA.
AC   P18520;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
```

```
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   INTERMEDIATE FILAMENT PROTEIN ON3.
OS   CARASSIUS AURATUS (GOLDFISH).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=GLIAL CELL;
RX   MEDLINE; 90180470.
RA   GIORDANO S., GLASGOW E., TESSER P., SCHECHTER N.;
RL   NEURON 2:1507-1516(1989).
CC   -!- FUNCTION: ONE OF THE NONNEURONAL PREDOMINANT INTERMEDIATE
CC       FILAMENT PROTEINS OF THE VISUAL PATHWAY.
CC   -!- SIMILARITY: SIGNIFICANT HOMOLOGY WITH MAMMALIAN TYPE II KERATIN
CC       K8.
DR   EMBL; M87773; -; NOT_ANNOTATED_CDS.
DR   PIR; JS0291; JS0291.
DR   PROSITE; PS00226; IF; 1.
KW   INTERMEDIATE FILAMENT; COILED COIL; HEPTAD REPEAT PATTERN; VISION.
FT   DOMAIN        1    108       HEAD.
FT   DOMAIN      109    420       ROD.
FT   DOMAIN      421    520       TAIL.
FT   DOMAIN      109    144       COIL 1A.
FT   DOMAIN      145    157       LINKER 1.
FT   DOMAIN      158    253       COIL 1B.
FT   DOMAIN      254    273       LINKER 12.
FT   DOMAIN      274    420       COIL 2.
SQ   SEQUENCE   520 AA;  57789 MW;  3E8CCC03 CRC32;

  Query Match             4.6%;  Score 96;  DB 1;  Length 520;
  Best Local Similarity  32.2%;  Pred. No. 1.66e-01;
  Matches    19;  Conservative   14;  Mismatches  24;  Indels   2;  Gaps   2;

Db    194 KNKYEDEIN-KRTECENDFVLIKKDVDEAYMNKVELEAKLESLSDEINFLRQIFEEEIR 251
          ||| :|: : : | | ||| |  |  : :|:|: :    ||: |  |  |  : | :::
Qy    156 KNKNDDDNDIMETLCKNDFALKIKVKEITYINR-DTKIILETKSKTIYKLNGVSERDLK 213


RESULT   11
ID   UD13_RAT       STANDARD;      PRT;   531 AA.
AC   Q64637;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   UDP-GLUCURONOSYLTRANSFERASE 1-3 PRECURSOR, MICROSOMAL (EC 2.4.1.17)
DE   (UDPGT) (UGT1*3) (UGT1-03) (UGT1.3) (UGT1A3) (B3).
GN   UGT1.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE OF 1-286 FROM N.A.
RC   STRAIN=WISTAR;
RX   MEDLINE; 95332265.
RA   EMI Y., IKUSHIRO S.I., IYANAGI T.;
RL   J. BIOCHEM. 117:392-399(1995).
RN   [2]
RP   SEQUENCE OF 287-531 FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 90274676.
RA   SATO H., KOIWAI O., TANABE K., KASHIWAMATA S.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 169:260-264(1990).
CC   -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC       SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC       ENDOGENOUS COMPOUNDS.
CC   -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC       BETA-D-GLUCORONOSIDE.
CC   -!- SUBCELLULAR LOCATION: MICROSOMAL.
CC   -!- ALTERNATIVE PRODUCTS: THE UGT1 GENE ENCODES FOR MANY DIFFERENT
CC       UDPGT ISOZYMES WHICH HAVE A DIFFERENT N-TERMINAL DOMAIN AND A
CC       COMMON C-TERMINAL DOMAIN OF 245 RESIDUES.
CC   -!- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
DR   EMBL; D38067; G940422; -.
DR   EMBL; M34007; G207579; ALT_TERM.
DR   PROSITE; PS00375; UDPGT; 1.
KW   TRANSFERASE; GLYCOSYLTRANSFERASE; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL;
KW   MULTIGENE FAMILY; MICROSOME; ALTERNATIVE SPLICING.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    531       UDP-GLUCURONOSYLTRANSFERASE 1-3.
FT   TRANSMEM    489    505       POTENTIAL.
FT   CARBOHYD    116    116       POTENTIAL.
FT   CARBOHYD    139    139       POTENTIAL.
FT   CARBOHYD    293    293       POTENTIAL.
FT   CARBOHYD    431    431       POTENTIAL.
SQ   SEQUENCE   531 AA;  60140 MW;  C3B2A846 CRC32;

  Query Match             4.5%;  Score 94;  DB 1;  Length 531;
  Best Local Similarity  35.7%;  Pred. No. 3.32e-01;
  Matches    15;  Conservative   15;  Mismatches   7;  Indels   5;  Gaps   5;

Db    214 PVP-WMYLCH-VNYGSLARLASDLLQREVSVVEILRHASVWL 253
          |:|  : ||| ::| :: ||: :|| :| :: |:| :|: |:
Qy     22 PIPANLQLCHGIEYQNM-RLP-NLLGHE-TMKEVLEQAGAWI 60


RESULT   12
ID   YM81_YEAST     STANDARD;      PRT;  1562 AA.
AC   Q04781; Q04029;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 180.2 KD PROTEIN IN FAA4-HOR7 INTERGENIC REGION.
GN   YMR247C OR YM9408.09C OR YM9920.01C.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE OF 1-956 FROM N.A.
RC   STRAIN=S288C / AB972;
RA   HUNT S., BOWMAN S., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL   SUBMITTED (MAR-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE OF 950-1562 FROM N.A.
RC   STRAIN=S288C / AB972;
RA   GENTLES S., BOWMAN S., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL   SUBMITTED (MAR-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z48639; G732925; -.
DR   EMBL; Z48756; G736313; -.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   1562 AA;  180185 MW;  CEE45672 CRC32;

  Query Match             4.4%;  Score 92;  DB 1;  Length 1562;
  Best Local Similarity  38.8%;  Pred. No. 6.58e-01;
  Matches    19;  Conservative    9;  Mismatches  21;  Indels   0;  Gaps   0;

Db   1397 KNEFDDINSKMDRLTSNDDALTIKLNNITNEVKASYLIDDQKLEISFKL 1445
          ||  || |  |: |  || || ||:::||   : : :| : |    :||
Qy    156 KNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKL 204


RESULT   13
ID   GYRB_ACIS3     STANDARD;      PRT;   216 AA.
AC   Q44273; Q60168;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   DNA GYRASE SUBUNIT B (EC 5.99.1.3) (FRAGMENTS).
GN   GYRB.
OS   ACINETOBACTER SP. (STRAIN ATCC 17924 / CIP 70.12).
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   NEISSERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97088999.
RA   YAMAMOTO S., HARAYAMA S.;
RL   INT. J. SYST. BACTERIOL. 46:506-511(1996).
CC   -!- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
```

```
CC       STRANDED DNA IN AN ATP-DEPENDANT MANNER AND ALSO CATALYZES THE
CC       INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC       DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC   -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
CC       OF DOUBLE-STRANDED DNA.
CC   -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC       BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC       ENZYME FORMS AN A2B2 TETRAMER.
DR   EMBL; D73436; G1322131; -.
DR   EMBL; D73421; G1322101; -.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   TOPOISOMERASE; ISOMERASE; ATP-BINDING; ANTIBIOTIC RESISTANCE.
FT   NON_TER       1      1
FT   NON_CONS    116    117
FT   NON_TER     216    216
SQ   SEQUENCE   216 AA;  23775 MW;  E572EEE0 CRC32;

  Query Match             4.3%;  Score 90;  DB 1;  Length 216;
  Best Local Similarity  26.3%;  Pred. No. 1.29e+00;
  Matches    20;  Conservative   23;  Mismatches  26;  Indels   7;  Gaps   7;

Db      79 LARRLRELSFLNAGVRIVLRDERVALEHIFDLEVGLSEKSAL-DIAGLPGKLA-DCQEKD 136
           |  :::|::::|    :|:| : : :   |: |: |:||:      :  |   |   |:|::
Qy     176 LKIKVKEITYINRDTKIIL-ETK-S-KTIYKLN-GVSERDLKKSVLWLKDSLQCTCEEMN 231

Db     137 PALSELYLVEGDSAGG 152
             ::  |||  |:  ||
Qy     232 -DINAPYLVMGQKQGG 246


RESULT   14
ID   VIRF_SHIDY     STANDARD;      PRT;   262 AA.
AC   Q04248;
DT   01-JUN-1994 (REL. 29, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   VIRULENCE REGULON TRANSCRIPTIONAL ACTIVATOR VIRF.
GN   VIRF.
OS   SHIGELLA DYSENTERIAE, SHIGELLA FLEXNERI, AND SHIGELLA SONNEI.
OG   PLASMID 210 KB INVASION.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.DYSENTERIAE; STRAIN=CG097;
RA   YAO R., REDDY L.V., PALCHAUDHURI S.;
RL   SUBMITTED (MAR-1991) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.FLEXNERI; STRAIN=2A;
RX   MEDLINE; 87032409.
RA   SAKAI T., SASAKAWA C., MAKINO S., YOSHIKAWA M.;
RL   INFECT. IMMUN. 54:395-402(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.SONNEI;
RX   MEDLINE; 89212881.
RA   KATO J.I., ITO K.I., NAKAMURA A., WATANABE H.;
RL   INFECT. IMMUN. 57:1391-1398(1989).
RN   [4]
RP   SIMILARITY TO ARAC FAMILY.
RX   MEDLINE; 92326642.
RA   DORMAN C.J.;
RL   MOL. MICROBIOL. 6:1575-1575(1992).
CC   -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR OF THE VIRB GENE WHICH IS
CC       ITSELF AN ACTIVATOR OF THE IPAABCD VIRULENCE REGULON.
CC   -!- SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL
CC       REGULATORS.
DR   EMBL; X58464; G46954; -.
DR   EMBL; M29172; G152801; -.
DR   EMBL; X16661; G47067; -.
DR   PIR; S14646; S14646.
DR   PIR; A47605; A47605.
DR   PIR; A60105; A60105.
DR   PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
DR   PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW   PLASMID; VIRULENCE; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING.
FT   DNA_BIND    177    196       H-T-H MOTIF (BY SIMILARITY).
SQ   SEQUENCE   262 AA;  30549 MW;  11339B38 CRC32;

  Query Match             4.3%;  Score 89;  DB 1;  Length 262;
  Best Local Similarity  25.5%;  Pred. No. 1.79e+00;
  Matches    13;  Conservative   21;  Mismatches  15;  Indels   2;  Gaps   2;

Db      37 TLTIDEGQIAFIERNIQINVSIKKSDSINPFEIISLDRNLLLSIIRIMEPI 87
           :| |   :|::|:|:  | :  | | :|  :: :| :|:|  |::::::::
Qy     175 ALKIKVKEITYINRDTKIILETK-SKTIYKLNGVS-ERDLKKSVLWLKDSL 223


RESULT   15
ID   KSGA_MYCPN     STANDARD;      PRT;   263 AA.
AC   P75113;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   DIMETHYLADENOSINE TRANSFERASE (EC 2.1.1.-) (S-ADENOSYLMETHIONINE-6-N',
DE   N'-ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RRNA DIMETHYLASE) (HIGH
DE   LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN
DE   DIMETHYLTRANSFERASE).
GN   KSGA.
OS   MYCOPLASMA PNEUMONIAE.
OC   PROKARYOTA; TENERICUTES; MOLLICUTES; MYCOPLASMA; MYCOPLASMATALES;
OC   MYCOPLASMATACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE; 97105885.
RA   HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
RA   HERRMANN R.;
RL   NUCLEIC ACIDS RES. 24:4420-4449(1996).
CC   -!- FUNCTION: SPECIFICALLY DIMETHYLATES TWO ADJACENT ADENOSINES IN THE
CC       LOOP OF A CONSERVED HAIRPIN NEAR THE 3'-END OF 16S RRNA IN THE 30S
CC       PARTICLE. ITS INACTIVATION LEADS TO KASUGAMYCIN RESISTANCE
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: STRONG, TO OTHER BACTERIAL KSGA.
CC   -!- SIMILARITY: BELONGS TO THE RRNA ADENINE N-6-METHYLTRANSFERASE
CC       FAMILY.
DR   EMBL; AE000017; G1673824; -.
DR   PROSITE; PS01131; RRNA_A_DIMETH; 1.
KW   MRNA PROCESSING; TRANSFERASE; METHYLTRANSFERASE;
KW   ANTIBIOTIC RESISTANCE.
SQ   SEQUENCE   263 AA;  29794 MW;  357AA00D CRC32;

  Query Match             4.3%;  Score 90;  DB 1;  Length 263;
  Best Local Similarity  37.0%;  Pred. No. 1.29e+00;
  Matches    20;  Conservative   11;  Mismatches  20;  Indels   3;  Gaps   3;

Db      52 KLQLPYHGIEL-DKRLAEYLLVNEILTEEQLTIGDALKQNLDQYFPDTIPLLCG 104
           :|||  ||||  ::||:: || :| : |     |  :   : |  |||  :||:
Qy      26 NLQL-CHGIEYQNMRLPN-LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCS 77


Search completed: Thu Oct 22 15:57:09 1998
Job time : 47 secs.
```



MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Oct 22 15:54:43 1998; MasPar time 9.36 Seconds
509.723 Million cell updates/sec

Tabular output not generated.

Title: >US-08-848-439-2
Description: (1-295) from US08848439.pep
Perfect Score: 2234
Sequence: 1 MLQGPGSLLLLFLASHCCLG.....WQKGQREFKRISRSIRKLQC 295

Scoring table: PAM 150
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 33.444; Variance 132.088; scale 0.253

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	396	17.7	572	27	W31273	Mouse frizzled-7 prot	2.72e-28
2	394	17.6	685	27	W31274	Mouse frizzled-8 prot	4.35e-28
3	374	16.7	585	27	W31271	Human frizzled-5 prot	4.69e-26
4	369	16.5	537	27	W31270	Mouse frizzled-4 prot	1.51e-25
5	357	16.0	525	27	W31269	Caenorhabditis frizzl	2.46e-24
6	355	15.9	694	27	W31267	Drosophila frizzled-2	3.92e-24
7	317	14.2	666	27	W31268	Mouse frizzled-3 prot	2.56e-20
8	284	12.7	709	27	W31272	Mouse frizzled-6 prot	4.86e-17
9	148	6.6	581	14	R74187	Chick p75.	3.74e-04
10	124	5.6	605	14	R74186	Chick p78.	4.55e-02
11	121	5.4	604	28	W35946	Human netrin-1.	8.16e-02
12	105	4.7	529	14	R74188	Mouse p78.	1.70e+00
13	100	4.5	732	25	W26642	Human RECK cancer-inh	4.26e+00
14	99	4.4	804	29	W55368	H. pylori ORF 09ap114	5.11e+00
15	96	4.3	868	25	W26610	Rat muscle-specific k	8.79e+00
16	96	4.3	868	25	W26507	Rat Dmk receptor.	8.79e+00
17	95	4.3	869	25	W26611	Human muscle-specific	1.05e+01
18	95	4.3	869	25	W26506	Human Dmk receptor.	1.05e+01
19	97	4.3	15281	9	R44929	T. niveum Cyclosporin	7.34e+00
20	94	4.2	467	17	R84094	Nsk2 receptor with pu	1.26e+01
21	93	4.2	478	16	R92718	Mouse muscle-localize	1.50e+01
22	93	4.2	530	16	R92715	Mouse muscle-localise	1.50e+01
23	93	4.2	860	16	R92716	Mouse muscle-localise	1.50e+01
24	93	4.2	868	16	R92717	Mouse muscle-localise	1.50e+01
25	87	3.9	343	6	R28887	pRi8 fusion protein.	4.28e+01
26	88	3.9	475	17	R94982	Nsk2 extracellular do	3.60e+01
27	87	3.9	530	6	R28888	Ri paraneoplastic ant	4.28e+01
28	87	3.9	550	13	R72830	Human TIABP2.	4.28e+01
29	88	3.9	580	29	W46758	Amino acid sequence o	3.60e+01
30	88	3.9	580	24	W25670	hNET sequence encoded	3.60e+01
31	88	3.9	740	19	R95911	Insulin-stimulated pr	3.60e+01
32	88	3.9	740	19	R95912	Mutant insulin-stimul	3.60e+01
33	88	3.9	863	17	R84088	Nsk2 receptor with al	3.60e+01
34	88	3.9	871	17	R84087	Nsk2 receptor.	3.60e+01
35	88	3.9	873	17	R84092	Nsk2 receptor with al	3.60e+01
36	88	3.9	881	17	R84091	Nsk2 receptor with al	3.60e+01
37	88	3.9	2516	20	W01875	Neuronal invertebrate	3.60e+01
38	88	3.9	2516	20	W01884	Invertebrate calcium	3.60e+01
39	87	3.9	2695	20	W07632	Human type I inositol	4.28e+01
40	87	3.9	2713	19	W00168	Human T cell inositol	4.28e+01
41	87	3.9	2749	3	R13887	Inositol-3-phosphate	4.28e+01
42	84	3.8	461	7	R37270	ODC.	7.15e+01
43	83	3.7	202	16	R77524	NK-Lysine full length	8.47e+01
44	82	3.7	311	8	R43682	Lipopolysaccharide in	1.00e+02
45	83	3.7	783	24	W27113	Rat spleen pro-hormon	8.47e+01

ALIGNMENTS

RESULT 1
ID W31273 standard; Protein; 572 AA.
AC W31273;
DT 27-APR-1998 (first entry)
DE Mouse frizzled-7 protein Mfz7 (Wnt receptor).
KW Wnt receptor; mouse frizzled-7 protein; Mfz7 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Mus musculus.
PN WO9739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV JOHNS HOPKINS.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR WPI; 97-526631/48.
DR N-PSDB; T89891.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 44-46; 61pp; English.
CC This protein comprises the mouse transmembrane receptor,
CC frizzled-7 (Mfz7), encoded by the Mfz7 gene (see T89891). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 572 AA;

Query Match 17.7%; Score 396; DB 27; Length 572;
Best Local Similarity 40.7%; Pred. No. 2.72e-28;
Matches 48; Conservative 25; Mismatches 42; Indels 3; Gaps 3;

Db 51 pisiplctdiaynqtilpnllghtnqedaglevhqfyplvkvqcspelrfflcsmyapvc 110

```
Qy      44 PANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVC 103

Db     111 -tvldqaippcrslcerarqgcealmnkfgfqwperlrcenfpvhgageicvgqntsd 167
             ||::| ||:||| : :: | ::|: ||| ||: | |: || :   ::|:   :||
Qy     104 LDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QDN-DLCIPLASSD 159


RESULT    2
ID    W31274 standard; Protein; 685 AA.
AC    W31274;
DT    27-APR-1998  (first entry)
DE    Mouse frizzled-8 protein Mfz8 (Wnt receptor).
KW    Wnt receptor; mouse frizzled-8 protein; Mfz8 gene;
KW    signal transduction; cancer; cell growth; cell proliferation.
OS    Mus musculus.
PN    WO9739357-A1.
PD    23-OCT-1997.
PF    11-APR-1997; U06049.
PR    12-APR-1996; US-015307.
PA    (UYJO ) UNIV JOHNS HOPKINS.
PA    (STRD ) UNIV LELAND STANFORD JUNIOR.
PI    Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI    Nusse R, Samos CH, Wangy;
DR    WPI; 97-526631/48.
DR    N-PSDB; T89892.
PT    Identification of Wnt receptor binding modulators - useful for
PT    treatment of cancer and growth, development or proliferation related
PT    disorders
PS    Disclosure; Page 48-50; 61pp; English.
CC    This protein comprises the mouse transmembrane receptor,
CC    frizzled-8 (Mfz8), encoded by the Mfz8 gene (see T89892).  It is
CC    an example of a Wnt receptor.  Other novel frizzled family members
CC    have been identified in human, mouse and Caenorhabditis elegans
CC    (see W31268-74) and are considered also to be Wnt receptors.  Wnt
CC    receptors can be used in a novel, claimed method of screening for
CC    compounds which modulate the binding of a Wnt polypeptide (secreted
CC    proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC    Wnt is involved in (mammary) cancer and other processes involving
CC    growth, development and proliferation (both normal and abnormal).
CC    Modulators identified by the claimed method are useful for
CC    treatment of diseases related to these conditions.
SQ    Sequence   685 AA;

  Query Match           17.6%;  Score 394;  DB 27;  Length 685;
  Best Local Similarity  43.8%;  Pred. No. 4.35e-28;
  Matches   56;  Conservative   23;  Mismatches  43;  Indels   6;  Gaps   6;

Db      40 vplckgigynytympnqfnhdtqdeaglevhqfw-plveiqcspdlkfflcsmytpicle 98
           : || || |    :|| : |:|  |  ||    | |||  || || | ||||:::|:||:
Qy      47 LQLCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLD 105

Db      99 dykkplppcrsvcerakagcaplmrqygfawpdrmrcdrlpeqgnpdtlcmdynrtd-lt 157
           |    : ||:|:| : |  |||:|  :||:||| : |||:| | | | ||:    :| |
Qy     106 DLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QDN-D-LCIPLASSDHLL 162

Db     158 taapsppr 165
            |:  :|:
Qy     163 PATEEAPK 170


RESULT    3
ID    W31271 standard; Protein; 585 AA.
AC    W31271;
DT    27-APR-1998  (first entry)
DE    Human frizzled-5 protein Mfz5 (Wnt receptor).
KW    Wnt receptor; human frizzled-5 protein; Hfz5 gene;
KW    signal transduction; cancer; cell growth; cell proliferation.
OS    Homo sapiens.
PN    WO9739357-A1.
PD    23-OCT-1997.
PF    11-APR-1997; U06049.
PR    12-APR-1996; US-015307.
PA    (UYJO ) UNIV JOHNS HOPKINS.
PA    (STRD ) UNIV LELAND STANFORD JUNIOR.
PI    Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI    Nusse R, Samos CH, Wangy;
DR    WPI; 97-526631/48.
DR    N-PSDB; T89889.
PT    Identification of Wnt receptor binding modulators - useful for
PT    treatment of cancer and growth, development or proliferation related
PT    disorders
PS    Disclosure; Page 37-39; 61pp; English.
CC    This protein comprises the human transmembrane receptor,
CC    frizzled-5 (Hfz5), encoded by the Hfz5 gene (see T89889).  It is
CC    an example of a Wnt receptor.  Other novel frizzled family members
CC    have been identified in human, mouse and Caenorhabditis elegans
CC    (see W31268-74) and are considered also to be Wnt receptors.  Wnt
CC    receptors can be used in a novel, claimed method of screening for
CC    compounds which modulate the binding of a Wnt polypeptide (secreted
CC    proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC    Wnt is involved in (mammary) cancer and other processes involving
CC    growth, development and proliferation (both normal and abnormal).
CC    Modulators identified by the claimed method are useful for
CC    treatment of diseases related to these conditions.
SQ    Sequence   585 AA;

  Query Match           16.7%;  Score 374;  DB 27;  Length 585;
  Best Local Similarity  38.8%;  Pred. No. 4.69e-26;
  Matches   47;  Conservative   29;  Mismatches  41;  Indels   4;  Gaps   4;

Db      20 lvgraaaaskapvcqei-t-vpmcrgigynlthmpnqfnhdtqdeaglevhqfw-plvei 76
           | |::  : | : |  | : : :|:|| |   ::|| : |:|  |  ||    | |||
Qy      27 LFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMK 85

Db      77 qcspdlrfflctmytpiclpdyhkplppcrsvcerakagcsplmrqygfawpermscdrl 136
           || || : |||:::|:|| |    : ||:|:| : |  |:|:|  :||:||: : |||:
Qy      86 QCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRF 145

Db     137 p 137
           |
Qy     146 P 146


RESULT    4
ID    W31270 standard; Protein; 537 AA.
AC    W31270;
DT    27-APR-1998  (first entry)
DE    Mouse frizzled-4 protein Mfz4 (Wnt receptor).
KW    Wnt receptor; mouse frizzled-4 protein; Mfz4 gene;
KW    signal transduction; cancer; cell growth; cell proliferation.
OS    Mus musculus.
PN    WO9739357-A1.
PD    23-OCT-1997.
PF    11-APR-1997; U06049.
PR    12-APR-1996; US-015307.
PA    (UYJO ) UNIV JOHNS HOPKINS.
PA    (STRD ) UNIV LELAND STANFORD JUNIOR.
PI    Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI    Nusse R, Samos CH, Wangy;
DR    WPI; 97-526631/48.
DR    N-PSDB; T89888.
PT    Identification of Wnt receptor binding modulators - useful for
PT    treatment of cancer and growth, development or proliferation related
PT    disorders
PS    Disclosure; Page 34-35; 61pp; English.
CC    This protein comprises the mouse transmembrane receptor,
CC    frizzled-4 (Mfz4), encoded by the Mfz4 gene (see T89888).  It is
CC    an example of a Wnt receptor.  Other novel frizzled family members
CC    have been identified in human, mouse and Caenorhabditis elegans
CC    (see W31268-74) and are considered also to be Wnt receptors.  Wnt
CC    receptors can be used in a novel, claimed method of screening for
CC    compounds which modulate the binding of a Wnt polypeptide (secreted
CC    proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC    Wnt is involved in (mammary) cancer and other processes involving
CC    growth, development and proliferation (both normal and abnormal).
CC    Modulators identified by the claimed method are useful for
```

```
CC   treatment of diseases related to these conditions.
SQ   Sequence   537 AA;

  Query Match           16.5%;  Score 369;  DB 27;  Length 537;
  Best Local Similarity  38.3%;  Pred. No. 1.51e-25;
  Matches    41;  Conservative   26;  Mismatches  40;  Indels   0;  Gaps   0;

Db     47 piriamcqnlgynvtkmpnlvghelqtdaelqlttftpliqygcssqlqfflcsvyvpmc 106
          |  : :|: : |   ::|||:|||   :   |  :  ||:   | ::   ||||:: |:|
Qy     44 PANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVC 103

Db    107 tekinipigpcggmclsvkrrcepvlrefgfawpdtlncskfppqnd 153
           : ::  | || ::|: || || ||:  |||:||| |:| :|| :||
Qy    104 LDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDND 150


RESULT    5
ID   W31269 standard; Protein; 525 AA.
AC   W31269;
DT   27-APR-1998  (first entry)
DE   Caenorhabditis frizzled-1 protein Cfz1 (Wnt receptor).
KW   Wnt receptor; Caenorhabditis frizzled-1 protein; Cfz1 gene;
KW   signal transduction; cancer; cell growth; cell proliferation.
OS   Caenorhabditis elegans.
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89887.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Disclosure; Page 30-31; 61pp; English.
CC   This protein comprises the Caenorhabditis putative transmembrane
CC   receptor, frizzled-1, encoded by the Cfz1 gene (see T89887).  It is
CC   an example of a Wnt receptor.  Other novel frizzled family members
CC   have been identified in human, mouse and Caenorhabditis elegans
CC   (see W31268-74) and are considered also to be Wnt receptors.  Wnt
CC   receptors can be used in a novel, claimed method of screening for
CC   compounds which modulate the binding of a Wnt polypeptide (secreted
CC   proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC   Wnt is involved in (mammary) cancer and other processes involving
CC   growth, development and proliferation (both normal and abnormal).
CC   Modulators identified by the claimed method are useful for
CC   treatment of diseases related to these conditions.
SQ   Sequence   525 AA;

  Query Match           16.0%;  Score 357;  DB 27;  Length 525;
  Best Local Similarity  31.8%;  Pred. No. 2.46e-24;
  Matches    54;  Conservative   51;  Mismatches  57;  Indels   8;  Gaps   8;

Db     32 emcndlpynltsfpnlvdeeswkdasesiltykpllsvvcseqlkfflcsvyfpmcnekl 91
          ::|: : |    :|||:  |: |:  |   :  ||:   |  : | ||||:: |:| : |
Qy     48 QLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDL 107

Db     92 anpigpcrplclsvqekclpvlesfgfkwpdvircdkfplennrekmcmk-gpneqg-ai 149
           : | ||::||: | ::| ||: :||| |||:: ||:|| ::| : :|:  ::::  :
Qy    108 DETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QDN-D-LCIPLASSDHLLPA 164

Db    150 qderakf-aa-keseddgndrvediqrevdrlngkcpqdevflnrsskcv 197
           :| :|   | |: :|| || :| : ::   |: |   :  ::|| :| :
Qy    165 TEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDTKII 213


RESULT    6
ID   W31267 standard; Protein; 694 AA.
AC   W31267;
DT   27-APR-1998  (first entry)
DE   Drosophila frizzled-2 protein (Wnt receptor).
KW   Wnt receptor; Drosophila frizzled-2 protein; Dfz2 gene;
KW   wingless receptor; Wg receptor; signal transduction; cancer;
KW   cell growth; cell proliferation.
OS   Drosophila melanogaster.
FH   Key             Location/Qualifiers
FT   Misc_difference 268
FT                   /note= "encoded by CAC"
FT   Misc_difference 269
FT                   /note= "encoded by TGG"
FT   Misc_difference 348
FT                   /note= "encoded by TA (apparent 1 nucleotide
FT                    deletion of codon)"
FT   Misc_difference 488
FT                   /note= "encoded by TTA"
FT   Misc_difference 632
FT                   /note= "encoded by CTG"
FT   Misc_difference 633
FT                   /note= "encoded by GCG"
FT   Misc_difference 671
FT                   /note= "encoded by CG (apparent 1 nucleotide
FT                    deletion of codon"
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89885.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Claim 2; Page 23-25; 61pp; English.
CC   This protein comprises the Drosophila frizzled-2 protein encoded by
CC   the Dfz2 gene (see T89885).  It is a receptor for wingless (Wg),
CC   acting as a signal transducing molecule, and is an example of a Wnt
CC   receptor (WntR).  Other novel frizzled family members have been
CC   identified in human, mouse and Caenorhabditis elegans (see
CC   W31268-74) and are considered also to be Wnt receptors.  Wnt
CC   receptors can be used in a novel, claimed method of screening for
CC   compounds which modulate the binding of a Wnt polypeptide (secreted
CC   proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC   Wnt is involved in (mammary) cancer and other processes involving
CC   growth, development and proliferation (both normal and abnormal).
CC   Modulators identified by the claimed method are useful for
CC   treatment of diseases related to these conditions.
SQ   Sequence   694 AA;

  Query Match           15.9%;  Score 355;  DB 27;  Length 694;
  Best Local Similarity  36.2%;  Pred. No. 3.92e-24;
  Matches    47;  Conservative   32;  Mismatches  45;  Indels   6;  Gaps   6;

Db     51 ygvpaip-kdpnlrceeitipmcrgigynmtsfpnemnhetqdeaglevhqfw-plveik 108
          :| | :: | :| :     : :|:|| |    :|| : |||  |  ||    | |||
Qy     28 FGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQ 86

Db    109 cspdlkfflcsmytpicledyhkplpvcrsvcerarsgcapimqqysfewpermacehlp 168
          | || | ||||:::|:||:|    :  |:|:| : :  |||:|  ::| ||: : |:::|
Qy     87 CHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP 146

Db    169 lhgdpdnlcm 178
           : | ::||:
Qy    147 -Q-D-NDLCI 153


RESULT    7
ID   W31268 standard; Protein; 666 AA.
AC   W31268;
DT   27-APR-1998  (first entry)
DE   Mouse frizzled-3 protein Mfz3 (Wnt receptor).
```

```
KW   Wnt receptor; mouse frizzled-3 protein; Mfz3 gene;
KW   signal transduction; cancer; cell growth; cell proliferation.
OS   Mus musculus.
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89886.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Disclosure; Page 27-29; 61pp; English.
CC   This protein comprises the mouse frizzled-3 protein encoded by
CC   the Mfz3 gene (see T89886), and is a Wnt receptor.  Other novel
CC   frizzled family members have been identified in human, mouse and
CC   Caenorhabditis elegans (see W31268-74) and are considered also to
CC   be Wnt receptors.  Wnt receptors can be used in a novel, claimed
CC   method of screening for compounds which modulate the binding of a
CC   Wnt polypeptide (secreted proteins involved in cell-to-cell
CC   signalling) to a Wnt receptor.  Wnt is involved in (mammary) cancer
CC   and other processes involving growth, development and proliferation
CC   (both normal and abnormal).  Modulators identified by the claimed
CC   method are useful for treatment of diseases related to these
CC   conditions.
SQ   Sequence   666 AA;

  Query Match           14.2%;  Score 317;  DB 27;  Length 666;
  Best Local Similarity  33.3%;  Pred. No. 2.56e-20;
  Matches    44;  Conservative   29;  Mismatches  54;  Indels   5;  Gaps   5;

Db      27 scepi-t-lrmcqdlpynttfmpnllnhydqqtaalame-pfhpmvnldcsrdfrpflca 83
           :| || : |::|: : |    :|||| | :     |    :  |:|  :|  | : |||:
Qy      39 NCKPIPANLQLCHGIEYQNMRLPNLLGH-ETMKEVLEQAGAWIPLVMKQCHPDTKKFLCS 97

Db      84 lyapicme-ygrvtlpcrrlcqraysecsklmemfgvpwpedmecsrfpdcdepyprlvd 142
           |:||:|::       ||: || :    |: :|  || |||: :|| |||: ::    |
Qy      98 LFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLAS 157

Db     143 lnlvgdptegap 154
            : :  :|| ||
Qy     158 SDHLLPATEEAP 169


RESULT    8
ID   W31272 standard; Protein; 709 AA.
AC   W31272;
DT   27-APR-1998  (first entry)
DE   Mouse frizzled-6 protein Mfz6 (Wnt receptor).
KW   Wnt receptor; mouse frizzled-6 protein; Mfz6 gene;
KW   signal transduction; cancer; cell growth; cell proliferation.
OS   Mus musculus.
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89890.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Disclosure; Page 41-42; 61pp; English.
CC   This protein comprises the mouse putative transmembrane receptor,
CC   frizzled-6 (Mfz6), encoded by the Mfz6 gene (see T89890).  It is
CC   an example of a Wnt receptor.  Other novel frizzled family members
CC   have been identified in human, mouse and Caenorhabditis elegans
CC   (see W31268-74) and are considered also to be Wnt receptors.  Wnt
CC   receptors can be used in a novel, claimed method of screening for
CC   compounds which modulate the binding of a Wnt polypeptide (secreted
CC   proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC   Wnt is involved in (mammary) cancer and other processes involving
CC   growth, development and proliferation (both normal and abnormal).
CC   Modulators identified by the claimed method are useful for
CC   treatment of diseases related to these conditions.
SQ   Sequence   709 AA;

  Query Match           12.7%;  Score 284;  DB 27;  Length 709;
  Best Local Similarity  33.6%;  Pred. No. 4.86e-17;
  Matches    41;  Conservative   26;  Mismatches  49;  Indels   6;  Gaps   6;

Db      24 cepitvp-r-cmkmty-nmtffpnlmghydqgiaavemghflhlanlecspniemflcqa 80
           | ||    : |  : | || : |||:||       : |  : |   :| |:  :|||
Qy      40 CKPIPANLQLCHGIEYQNMRL-PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSL 98

Db      81 fiptcteqih-vvlpcrklcekivsdckklmdtfgirwpeelecnrlphcddt-vpvtsh 138
           | | | :::   : ||: ||  :   |  :| :||: ||: |||:|:|: :|  :|::|
Qy      99 FAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASS 158

Db     139 ph 140
             |
Qy     159 DH 160


RESULT    9
ID   R74187 standard; Protein; 581 AA.
AC   R74187;
DT   24-DEC-1995  (first entry)
DE   Chick p75.
KW   Neural axon out-growth modulator; epidermal growth factor; EGF;
KW   netrin-2; p75; neurodegenerative disease; transgenic animal;
KW   gene therapy.
OS   Gallus sp.
FH   Key             Location/Qualifiers
FT   peptide         1..15
FT                   /label= Sig_peptide
FT   domain          16..260
FT                   /label= Domain-VI
FT                   /note= "domain VI is probably involved in protein-
FT                   protein complex formation"
FT   domain          261..429
FT                   /label= Domain-V
FT                   /note= "domain V includes EGF-like repeats"
FT   region          430..581
FT                   /note= "C-terminal region"
PN   WO9513367-A1.
PD   18-MAY-1995.
PF   08-NOV-1994; U12913.
PR   12-NOV-1993; US-152019.
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
PI   Dodd J,  Jessell T,  Kennedy T,  Placzek M,  Serafini T;
PI   Tessier-Lavigne M;
DR   WPI; 95-194086/25.
DR   N-PSDB; Q92367.
PT   Neural axon out-growth modulators derived from EGF-like repeats of
PT   netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
PT   increasing spinal axon out-growth or directing axon orientation
PS   Claim 1; Page 48-50; 58pp; English.
CC   An E10 chick brain cDNA library was screened with probes based on
CC   netrin-1 (p78) or netrin-2 (p75) sequences to isolate chick p78
CC   and p75 partial cDNA clones.  Full-length clones (given in
CC   Q92366-67, respectively) were subsequently obtd. by 3'RACE.  cDNA
CC   is expressed e.g. in COS or insect cells for recombinant p78 and
CC   p75 prodn., used to breed transgenic animals, or for gene therapy.
SQ   Sequence   581 AA;

  Query Match            6.6%;  Score 148;  DB 14;  Length 581;
  Best Local Similarity  27.0%;  Pred. No. 3.74e-04;
```

```
  Matches    30;  Conservative   28;  Mismatches  46;  Indels   7;  Gaps   6;

Db     431 ipainptslvtste-apadcdsyckpakgnykinmkkyckkdyvvqvnilemetvanwak 489
           ||  ::  |: :|| ||  |:: ||  : : :  |   ||:|: : ::: |:  :   :|
Qy     153 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTK 211

Db     490 ftinils--vykcrdervkr-gdnflwihlkdlsckcpki-qiskkylvmg 536
           : ::  |  :||      :    : ||:    | | |  : :|:  |||||
Qy     212 IILETKSKTIYKLNGVSERDLKKSVLWLK-DSLQCTCEEMNDINAPYLVMG 261


RESULT   10
ID   R74186 standard; Protein; 605 AA.
AC   R74186;
DT   24-DEC-1995  (first entry)
DE   Chick p78.
KW   Neural axon out-growth modulator; epidermal growth factor; EGF;
KW   netrin-1; p78; neurodegenerative disease; transgenic animal;
KW   gene therapy.
OS   Gallus sp.
FH   Key             Location/Qualifiers
FT   peptide         1..25
FT                   /label= Sig_peptide
FT   misc_difference 29
FT                   /note= "unspecified amino acid"
FT   misc_difference 38
FT                   /note= "unspecified amino acid"
FT   misc_difference 50
FT                   /note= "unspecified amino acid"
FT   domain          26..284
FT                   /label= Domain-VI
FT                   /note= "domain VI is probably involved in protein-
FT                   protein complex formation"
FT   domain          285..453
FT                   /label= Domain-V
FT                   /note= "domain V includes EGF-like repeats"
FT   region          454..605
FT                   /note= "C-terminal region"
PN   WO9513367-A1.
PD   18-MAY-1995.
PF   08-NOV-1994; U12913.
PR   12-NOV-1993; US-152019.
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
PI   Dodd J,  Jessell T,  Kennedy T,  Placzek M,  Serafini T;
PI   Tessier-Lavigne M;
DR   WPI; 95-194086/25.
DR   N-PSDB; Q92366.
PT   Neural axon out-growth modulators derived from EGF-like repeats of
PT   netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
PT   increasing spinal axon out-growth or directing axon orientation
PS   Claim 1; Page 44-46; 58pp; English.
CC   An E10 chick brain cDNA library was screened with probes based on
CC   netrin-1 (p78) or netrin-2 (p75) sequences to isolate chick p78
CC   and p75 partial cDNA clones.  Full-length clones (given in
CC   Q92366-67, respectively) were subsequently obtd. by 3'RACE.  cDNA
CC   is expressed e.g. in COS or insect cells for recombinant p78 and
CC   p75 prodn., used to breed transgenic animals, or for gene therapy.
SQ   Sequence   605 AA;

  Query Match             5.6%;  Score 124;  DB 14;  Length 605;
  Best Local Similarity  26.1%;  Pred. No. 4.55e-02;
  Matches    31;  Conservative   34;  Mismatches  41;  Indels  13;  Gaps  11;

Db     455 ipaappptaassteepad-cdsyckaskgklkinmkkyckkdyavqihilkaeknadw-w 512
           || |::    ::|||::  |:: ||  :   :  |   ||:|:|: | : |     ::
Qy     153 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDT 210

Db     513 kftvniis--vyk-qg-snr-lrrgdqtlwvhakdiackcpkvkpmkk-ylllgsteds 565
           |: ::  |  :||  | |:| |:::   ||:    : | |  :: ::  ||::|  : :
Qy     211 KIILETKSKTIYKLNGVSERDLKKS--VLWLKD-SLQCTCEEMNDINAPYLVMGQKQGG 266


RESULT   11
ID   W35946 standard; Protein; 604 AA.
AC   W35946;
DT   11-MAY-1998  (first entry)
DE   Human netrin-1.
KW   Netrin-1; neuron; growth; differentiation; morphology;
KW   neural disease; diagnosis; therapy; drug screening; human.
OS   Homo sapiens.
PN   WO9740064-A1.
PD   30-OCT-1997.
PF   19-APR-1997; U06452.
PR   19-APR-1996; US-635137.
PA   (EXEL-) EXELIXIS PHARM INC.
PA   (REGC ) UNIV CALIFORNIA.
PI   Kennedy T, Leonardo D, Serafini T, Shyian A, Swimmer C,
PI   Tessier-Lavigne M, Zhang Y;
DR   WPI; 97-535773/49.
DR   N-PSDB; T97129.
PT   Human netrin-1 protein and related nucleic acids - useful in
PT   modulating neuron growth and screening for compounds for diagnosis
PT   or treatment of diseases associated with undesirable growth
PS   Claim 1; Page 13-15; 22pp; English.
CC   This protein comprises human netrin-1 (see W35946), a protein
CC   which is involved in neural axon guidance, and which is especially
CC   useful in modulating neural axon outgrowth.  Its amino acid
CC   sequence was deduced from a cDNA clone (see T97129) isolated from a
CC   human foetal brain cDNA library.  Neuron growth, differentiation or
CC   morphology can be altered by contact with netrin (claimed).
CC   Isolated netrin-1 can also be used to screen chemical libraries
CC   for candidate drugs suitable for diagnosis or treatment of diseases
CC   associated with undesirable neural cell growth, by comparing
CC   binding to a netrin binding target with and without the presence of
CC   a prospective agent.  Agents that modulate the interaction may be
CC   useful as pharmaceutical lead compounds (claimed).
SQ   Sequence   604 AA;

  Query Match             5.4%;  Score 121;  DB 28;  Length 604;
  Best Local Similarity  24.8%;  Pred. No. 8.16e-02;
  Matches    29;  Conservative   30;  Mismatches  49;  Indels   9;  Gaps   8;

Db     454 ipvappttaassvee-pedcdsyckaskgklkinmkkyckkdyavqihilkadkagdw-w 511
           ||:|::    :: || |  |:: ||  :   :  |   ||:|:|: | : |     ::
Qy     153 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDT 210

Db     512 kftvniis--vykqgtsrirrgdqs-lwirsrdiackcpkikplkk-ylllgnaeds 564
           |: ::  |  :||      |    | ||::   : | |  :: ::  ||::|  : :
Qy     211 KIILETKSKTIYKLNGVSERDLKKSVLWLKD-SLQCTCEEMNDINAPYLVMGQKQGG 266


RESULT   12
ID   R74188 standard; Protein; 529 AA.
AC   R74188;
DT   24-DEC-1995  (first entry)
DE   Mouse p78.
KW   Neural axon out-growth modulator; epidermal growth factor; EGF;
KW   netrin-1; p78; neurodegenerative disease; transgenic animal;
KW   gene therapy.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   misc_difference 296..297
FT                   /note= "unidentified amino acids"
FT   misc_difference 300..301
FT                   /note= "unidentified amino acids"
FT   misc_difference 307..308
FT                   /note= "unidentified amino acids"
FT   misc_difference 311
FT                   /note= "unidentified amino acid"
FT   misc_difference 389
FT                   /note= "unidentified amino acid"
FT   misc_difference 393..394
FT                   /note= "unidentified amino acids"
FT   misc_difference 487
FT                   /note= "unidentified amino acid"
```

PN WO9513367-A1.
PD 18-MAY-1995.
PF 08-NOV-1994; U12913.
PR 12-NOV-1993; US-152019.
PA (REGC) UNIV CALIFORNIA.
PA (UYCO) UNIV COLUMBIA NEW YORK.
PI Dodd J, Jessell T, Kennedy T, Placzek M, Serafini T;
PI Tessier-Lavigne M;
DR WPI; 95-194086/25.
DR N-PSDB; Q92368.
PT Neural axon out-growth modulators derived from EGF-like repeats of
PT netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
PT increasing spinal axon out-growth or directing axon orientation
PS Claim 1; Page 51-52; 58pp; English.
CC Chick p75 and p78 sequences (given in R74186-87, respectively)
CC were used to identify conserved amino acid regions, which were
CC then used to design degenerate primers for the amplification
CC of a fragment of the mouse p78 cDNA (Q92368). This cDNA can
CC be expressed in host cells for recombinant p78 prodn, or used
CC to breed transgenic animals, or for gene therapy.
SQ Sequence 529 AA;

Query Match 4.7%; Score 105; DB 14; Length 529;
Best Local Similarity 23.5%; Pred. No. 1.70e+00;
Matches 24; Conservative 25; Mismatches 45; Indels 8; Gaps 7;

```
Db    388 sxveexxedcdsyckaskgklkmnmkkycrkdyavqihilkadkagdw-wkftvniis-- 444
          :  ||    |:: ||  :   :  |   |::|:|: | : |     ::  |: ::  |
Qy    163 PATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDTKIILETKSKT 220

Db    445 vykqgtsrirrgdqs-lwirsrdiackcpkikplkk-ylllg 484
          :||      |    | ||::   : | |  :: ::  ||::|
Qy    221 IYKLNGVSERDLKKSVLWLKD-SLQCTCEEMNDINAPYLVMG 261
```

RESULT 13
ID W26642 standard; Protein; 732 AA.
AC W26642;
DT 11-FEB-1998 (first entry)
DE Human RECK cancer-inhibiting protein.
KW RECK; reversion-inducing cysteine rich protein with Kazal motif;
KW human; cancer; gene therapy; diagnosis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 396..415
FT /label= Kazal_domain
PN WO9724439-A1.
PD 10-JUL-1997.
PF 24-DEC-1996; U20812.
PR 27-DEC-1995; JP-340469.
PA (AMGE-) AMGEN INC.
PA (KITA/) KITAYAMA H.
PA (NODA/) NODA M.
PA (TAKA/) TAKAHASHI C.
PA (SANY) SANKYO CO LTD.
PI Kitayama H, Noda M, Takahashi C;
DR WPI; 97-363675/33.
DR N-PSDB; T90508.
PT RECK gene and corresponding protein sequences - enables reversion of
PT cancer cells
PS Claim 7; Page 42-44; 53pp; English.
CC This protein sequence comprises the human RECK protein (reversion-
CC inducing cysteine rich protein with Kazal motif), which is capable
CC of transforming malignant cancer cells with an activated ras gene
CC into normal cells (reversion activity). Its amino acid sequence
CC was deduced from a cDNA clone (see T90508) obtained from human
CC fibroblast MRC-5 (ATCC CCL 171) cells. A claimed method for
CC treatment of cancer comprises contacting the cancer cells with a
CC RECK polypeptide. RECK may also be expressed using gene therapy
CC methods for in vivo treatment of cancer.
SQ Sequence 732 AA;

Query Match 4.5%; Score 100; DB 25; Length 732;
Best Local Similarity 30.5%; Pred. No. 4.26e+00;
Matches 25; Conservative 15; Mismatches 36; Indels 6; Gaps 5;

```
Db    158 ipvldikkcqpemwkaiacslqikpchsksrgsii-cksdcveilkkcgd-qnkfpedht 215
          || | :| |:|:  | : |||    |      :|  | | ||::  :|:   : |
Qy     80 IP-LVMKQCHPDT-KKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWP 137

Db    216 aesicellsptddlkncipldt 237
              |: ::  :||  ||||  :
Qy    138 DMLECDRFPQDNDL--CIPLAS 157
```

RESULT 14
ID W55368 standard; Protein; 804 AA.
AC W55368;
DT 17-JUN-1998 (first entry)
DE H. pylori ORF 09ap11406orf2 protein.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
OS Helicobacter pylori.
PN WO9737044-A1.
PD 09-OCT-1997.
PF 27-MAR-1997; U05223.
PR 06-DEC-1996; US-761318.
PR 29-MAR-1996; US-625811.
PR 02-APR-1996; US-758731.
PR 25-OCT-1996; US-736905.
PR 28-OCT-1996; US-738859.
PA (ASTR) ASTRA AB.
PI Alm RA, Smith D;
DR WPI; 97-503122/46.
DR N-PSDB; V24777.
PT Helicobacter pylori nucleic acid sequences and encoded
PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
PT infection and for diagnosis of H. pylori infection
PS Claim 14; Page 587-589; 1145pp; English.
CC This sequence is a H. pylori protein of unspecified function.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors. The
CC DNA and probes derived from it may be used for the identification of
CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
CC acid sequences complementary to the DNA act as antisense sequences and
CC can be used to prevent the translation of H. pylori mRNA. Antibodies
CC against the protein can be used in immunoassays to evaluate the abundance
CC and distribution of H. pylori-specific antigens. The genomic sequence of
CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
CC by mechanically shearing the bacterial DNA. The sequences were analysed
CC for ORF of at least 180 nucleotides, and the predicted coding regions
CC defined by computer evaluation. To identify likely H. pylori antigens for
CC vaccine development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or exported
CC membrane proteins. Having identified and determined the sequences of
CC interest, particular regions can be isolated from H. pylori by PCR
CC amplification for recombinant polypeptide production, e.g. in E. coli
CC hosts.
SQ Sequence 804 AA;

Query Match 4.4%; Score 99; DB 29; Length 804;
Best Local Similarity 31.5%; Pred. No. 5.11e+00;
Matches 17; Conservative 16; Mismatches 19; Indels 2; Gaps 2;

```
Db     78 vidsipkegftplenaffnitysmrqqqfilhikpkvmrrltlfsfdrdykkai 131
          ::::: |: |: |     :|||  |:  :||  |:| : :|   | :|| ||::
Qy    185 IMETLCKNDFA-LKIKVKEITYINRDTKIILETKSKTIYKLNGVS-ERDLKKSV 236
```

RESULT 15
ID W26610 standard; Protein; 868 AA.
AC W26610;
DT 27-JAN-1998 (first entry)
DE Rat muscle-specific kinase (MuSK).

KW Receptor tyrosine kinase; muscle specific kinase; MuSK; Dmk; rat;
KW ligand; agrin; diagnosis; therapy.
OS Rattus sp.
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= Sig_peptide
FT Domain 21..492
FT /label= Extracellular_domain
FT Domain 493..521
FT /label= Transmembrane_domain
FT Domain 522..868
FT /label= Intracellular_domain
PN WO9721811-A2.
PD 19-JUN-1997.
PF 13-DEC-1996; U20696.
PR 10-MAY-1996; US-644271.
PR 15-DEC-1995; US-008657.
PA (REGE-) REGENERON PHARM INC.
PI Bowen DC, Glass DJ, Valenzuela DM, Yancopoulos GD;
DR WPI; 97-332783/30.
DR N-PSDB; T90472.
PT Nucleotide sequences encoding human agrin and muscle specific kinase
PT and related receptor - used in diagnosis and treatment of disorder
PT with muscle atrophy
PS Example 1; Fig 1; 120pp; English.
CC This polypeptide comprise a novel rat receptor tyrosine kinase
CC designated muscle specific kinase (MuSK) that is expressed in
CC normal and denervated muscle. MuSK is alternatively referred to
CC Dmk for denervated muscle kinase. The amino acid sequence was
CC deduced from an isolated cDNA clone (see T90472). Human MuSK (see
CC W26611) has also been identified. Use of MuSK to generate anti-
CC MuSK antibodies and in the diagnosis of neurological or other
CC disorders is disclosed. Assay systems that may be used to detect
CC and/or measure ligands that bind the musk gene product are provided.
CC A claimed method of promoting the growth, differentiation or
CC survival of MuSK receptor-expressing cells involves administration
CC to the cell of agrin (see W26609). Such cells include muscle,
CC heart, spleen, ovary and retina cells, or cells genetically
CC engineered to express the MuSK receptor.
SQ Sequence 868 AA;

Query Match 4.3%; Score 96; DB 25; Length 868;
Best Local Similarity 29.7%; Pred. No. 8.79e+00;
Matches 19; Conservative 12; Mismatches 28; Indels 5; Gaps 5;

```
Db    345 eeaqelliht-awnelkavsplcrpaaeallcnhlfqecspgvlptpmpicreyclavke 403
          |   |:| :: ||  |  :   |:| :  :||: ||       |   :  |:  |: ||:
Qy     67 ETMKEVLEQAGAWIPL-VMKQ-CHPDTKKFLCS-LFAPVCLDDLDETIQPCHSLCVQVKD 123

Db    404 lfca 407
           ||
Qy    124 R-CA 126
```

Search completed: Thu Oct 22 15:56:03 1998
Job time : 80 secs.

FSEDGYRTVVQGTKKEGCTILFMML
NSQYFHLAAWAVPAVKTITILAMGQ
IGTSFLLAGFVSLFRIRTIMKHDGT
QAFREHWERSWVSQHCKSLAIPCPA
GKTLHSWRKFYTRLTNSRHGETTV"

ength 1923;

Indels 3; Gaps 2;

caaccttctgggccaca 322
||::| :| || ||:
AAYYTNYTNGGNCAYG 199

gctggtgaaggtgcagt 382
:| || | | :||:|
NYTNGTNATGAARCART 259

cgtgtgcaccg-tg--c 439
|| ||: | :| :
NGTNTGYYTNGAYGAYY 319

gcgccagggctgcgaag 499
| | ||:|
NAARGAYMGNTGYGCNC 379

ctgcgagcacttcccgc 559
||:|| : ||:|| |
RTGYGAYMGNTTYCCNC 439

OD 16-JUL-1993
e (frizzled) homologue

) osteosarcoma cDNA to

tazoa; Chordata;
athi; Myomorpha; Muridae;

ooks,M., Bradley,M.S.,
,C.D., Strewler,G.J. and

gene frizzled (fz) are

(frizzled) homologue"

/db_xref="PID:g310115"
/translation="MRARSALPRSALPRLLLPLLLLPAAGPAQFHGEKGISIPDHGFC
QPISIPLCTDIAYNQTIMPNLLGHTNQEDAGLEVHQFYPLVKVQCSPELRFFLCSMYA
PVCTVLEQAIPPCRSICERARQGCEALMNKFGFQWPERLRCEHFPRHGAEQICVGQNH
SEDGTPALLTTAPPSGLQPGAGGTPGGPGGGGAPPRYATLEHPFHCPRVLKVPSYLSY
KFLGERDCAAPCEPARPDGSMFFSHHHTRFARLWILTWSVLCCASTFFTVTTSLVAMQ
RFRYPERPIIFLSGCYTMVSVAYIAGFVLQERVVCNERFSEDGYRTVGQGTKKEGCTI
LFMMLYFFSMASSIWWVILSLTWFLAAGMKWGHAAIEANSQYFHLAAWAVPAVKTITI
LAMGQIDGDLLSGVCFVGLNRLDPLRGFVLAPLFVYLFIGTSFLLAGFVSLFRIRTIM
KHDGTKTEPLERLMVRIGVFSVLYTVPATIVIACYFYEQAFREHWERSWVSQHCKSLA
IPCPAHYTPRTSPDFTVYMIKYLMTLIVGITSGFWIWSGKTLHSWRKFYTRLTNSRHG
ETTV"
BASE COUNT 299 a 662 c 566 g 385 t
ORIGIN

Query Match 7.1%; Score 52; DB 90; Length 1912;
Best Local Similarity 39.3%; Pred. No. 2.64e-06;
Matches 118; Conservative 41; Mismatches 138; Indels 3; Gaps 2;

Db 224 tcccgctgtgcacggacatcgcctacaaccagaccatcatgcccaaccttcttgggcaca 283
| | ::| ||: | || | ||: | | | : | || ||::| :| || ||:
Qy 140 TNCARYTNTGYCAYGGNATNGARTAYCARAAYATGMGNYTNCCNAAYYTNYTNGGNCAYG 199

Db 284 cgaaccaagaggacgcgggcctggaggtgcatcaattctacccgctggtgaaggtgcagt 343
:| |:|| | : |:| | || :| || | | :||:|
Qy 200 ARACNATGAARGARGTNYTNGARCARGCNGGNGCNTGGATNCCNYTNGTNATGAARCART 259

Db 344 gctcgcccgagctgcgcttcttcctgtgctccatgtacgctccggtgtgcacgg-tg--c 400
|: || || ||::| ||::: | | :|| || || ||: | :| :
Qy 260 GYCAYCCNGAYACNAARAARTTYYTNTGYWSNYTNTTYGCNCCNGTNTGYYTNGAYGAYY 319

Db 401 tggagcaggccatcccgccgtgccgctccatctgcgaacgcgcgcgccaaggctgcgagg 460
| || |: | || | :|| ||:| ::: | ||:| | | | | ||:|
Qy 320 TNGAYGARACNATNCARCCNTGYCAYWSNYTNTGYGTNCARGTNAARGAYMGNTGYGCNC 379

Db 461 cgctcatgaacaagttcggcttccagtggcccgagcgcctccgctgcgagcatttcccgc 520
| | |||: ||:|| ||:| ||||| || :| ||:|| : ||:|| |
Qy 380 CNGTNATGWSNGCNTTYGGNTTYCCNTGYCCNGAYATGYTNGARTGYGAYMGNTTYCCNC 439

RESULT 12
ID CET23D8 standard; DNA; INV; 34576 BP.
AC Z81128;
NI e1008304
DT 21-OCT-1996 (Rel. 49, Created)
DT 01-MAR-1997 (Rel. 51, Last updated, Version 5)
DE Caenorhabditis elegans cosmid T23D8
KW CD63 antigen like; Drosophila tissue polarity protein like;
KW histone H2A; histone H4; Mouse bright protein like;
KW transporter protein.
OS Caenorhabditis elegans
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
OC Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP 1-34576
RA Wild A.;
RT ;
RL Submitted (21-OCT-1996) to the EMBL/GenBank/DDBJ databases.
RL Nematode Sequencing Project, Sanger Centre, Hinxton, Cambridge CB10
RL 1RQ, England and Department of Genetics, Washington University, St.
RL Louis, MO 63110, USA. E-mail: jes@sanger.ac.uk or
RL rw@nematode.wustl.edu
RN [2]

```
****************************************************************************
MPSRCH (TM)
****************************************************************************
```



```
MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Oct 22 15:50:45 1998;  MasPar time 9.06 Seconds
                                            817.121 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-848-439-2
Description:     (1-295) from US08848439.pep
Perfect Score:   2234
Sequence:        1 MLQGPGSLLLLFLASHCCLG..........WQKGQREFKRISRSIRKLQC 295

Scoring table:   PAM 150
                 Gap 11

Searched:        69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        swiss-prot35
                   1:swiss1

Statistics:      Mean 46.742;  Variance 71.562;  scale 0.653
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	332	14.9	581	1	FRIZ_DROME	FRIZZLED PROTEIN PRECU	9.66e-53
2	148	6.6	581	1	NET2_CHICK	NETRIN-2 PRECURSOR (FR	8.14e-11
3	124	5.6	606	1	NET1_CHICK	NETRIN-1 PRECURSOR.	3.06e-06
4	109	4.9	465	1	YQI8_CAEEL	HYPOTHETICAL 53.8 KD P	1.32e-03
5	105	4.7	500	1	CPD4_RAT	CYTOCHROME P450 IID4 (	6.13e-03
6	105	4.7	500	1	CPDI_RAT	CYTOCHROME P450 IID18	6.13e-03
7	101	4.5	1741	1	RPC1_GIALA	DNA-DIRECTED RNA POLYM	2.73e-02
8	99	4.4	724	1	KS61_MOUSE	RIBOSOMAL PROTEIN S6 K	5.67e-02
9	98	4.4	1032	1	VG07_BPT4	BASEPLATE STRUCTURAL P	8.13e-02
10	96	4.3	520	1	ION3_CARAU	INTERMEDIATE FILAMENT	1.66e-01
11	94	4.2	531	1	UD13_RAT	UDP-GLUCURONOSYLTRANSF	3.35e-01
12	92	4.1	1562	1	YM81_YEAST	HYPOTHETICAL 180.2 KD	6.68e-01
13	90	4.0	216	1	GYRB_ACIS3	DNA GYRASE SUBUNIT B (	1.31e+00
14	89	4.0	262	1	VIRF_SHIDY	VIRULENCE REGULON TRAN	1.83e+00
15	90	4.0	263	1	KSGA_MYCPN	DIMETHYLADENOSINE TRAN	1.31e+00
16	90	4.0	561	1	VATA_MAIZE	VACUOLAR ATP SYNTHASE	1.31e+00
17	89	4.0	752	1	KS6A_CHICK	RIBOSOMAL PROTEIN S6 K	1.83e+00
18	88	3.9	127	1	YBU9_YEAST	HYPOTHETICAL 14.8 KD P	2.55e+00
19	88	3.9	182	1	YLP2_CAEEL	HYPOTHETICAL 20.1 KD P	2.55e+00
20	88	3.9	443	1	AM2A_ORYSA	ALPHA-AMYLASE ISOZYME	2.55e+00
21	88	3.9	445	1	AMC2_ORYSA	ALPHA-AMYLASE ISOZYME	2.55e+00
22	87	3.9	446	1	YHE2_PSEAE	HYPOTHETICAL 50.0 KD P	3.53e+00
23	87	3.9	510	1	NOA1_HUMAN	ONCONEURAL VENTRAL ANT	3.53e+00
24	87	3.9	580	1	VATA_HORVU	VACUOLAR ATP SYNTHASE	3.53e+00
25	88	3.9	633	1	KS62_MOUSE	RIBOSOMAL PROTEIN S6 K	2.55e+00
26	88	3.9	740	1	KS62_HUMAN	RIBOSOMAL PROTEIN S6 K	2.55e+00
27	88	3.9	848	1	NFM_MOUSE	NEUROFILAMENT TRIPLET	2.55e+00
28	88	3.9	1102	1	YK64_CAEEL	HYPOTHETICAL 124.8 KD	2.55e+00
29	87	3.9	1201	1	COPA_YEAST	COATOMER ALPHA SUBUNIT	3.53e+00
30	87	3.9	1752	1	DESP_HUMAN	DESMOPLAKIN I AND II (	3.53e+00
31	87	3.9	2386	1	FINC_HUMAN	FIBRONECTIN PRECURSOR	3.53e+00
32	87	3.9	2749	1	IP3R_MOUSE	INOSITOL 1,4,5-TRISPHO	3.53e+00
33	87	3.9	2749	1	IP3R_RAT	INOSITOL 1,4,5-TRISPHO	3.53e+00
34	86	3.8	327	1	VMP_CAMVD	MOVEMENT PROTEIN (CELL	4.88e+00
35	86	3.8	450	1	DCOR_CHICK	ORNITHINE DECARBOXYLAS	4.88e+00
36	86	3.8	461	1	DCOR_BOVIN	ORNITHINE DECARBOXYLAS	4.88e+00
37	85	3.8	495	1	E1BL_ADE02	E1B PROTEIN, LARGE T-A	6.72e+00
38	86	3.8	556	1	FTHS_STRMU	FORMATE--TETRAHYDROFOL	4.88e+00
39	86	3.8	620	1	Y870_METJA	HYPOTHETICAL PROTEIN M	4.88e+00
40	86	3.8	725	1	VACB_MYCGE	VACB PROTEIN HOMOLOG.	4.88e+00
41	85	3.8	798	1	VP16_YEAST	VACUOLAR PROTEIN SORTI	6.72e+00
42	85	3.8	1701	1	MSP1_PLAFF	MEROZOITE SURFACE PROT	6.72e+00
43	85	3.8	1726	1	MSP1_PLAFC	MEROZOITE SURFACE PROT	6.72e+00
44	85	3.8	1726	1	MSP1_PLAFP	MEROZOITE SURFACE PROT	6.72e+00
45	85	3.8	2688	1	ZEP1_MOUSE	ZINC FINGER PROTEIN 40	6.72e+00

ALIGNMENTS

```
RESULT    1
ID   FRIZ_DROME     STANDARD;      PRT;   581 AA.
AC   P18537;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT   01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE   FRIZZLED PROTEIN PRECURSOR.
GN   FZ.
OS   DROSOPHILA MELANOGASTER (FRUIT FLY).
OC   EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN   [1]
RP   SEQUENCE FROM N.A. (CLASS I).
RX   MEDLINE; 89159415.
RA   VINSON C.R., CONOVER S., ADLER P.N.;
RL   NATURE 338:263-264(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (CLASSES I AND II).
RX   MEDLINE; 91060073.
RA   ADLER P.N., VINSON C., PARK W.J., CONOVER S., KLEIN L.;
RL   GENETICS 126:401-416(1990).
CC   -!- FUNCTION: REQUIRED TO COORDINATE THE CYTOSKELETONS OF EPIDERMAL
CC       CELLS TO PRODUCE A PARALLEL ARRAY OF CUTICULAR HAIRS AND BRISTLES.
CC       FZ IS REQUIRED FOR THE TRANSMISSION OF POLARITY INFORMATION, AND
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: CLASS I AND CLASS II PROTEINS DIFFER AT THE
CC       C-TERMINAL FROM AA 406 DUE TO USE OF ALTERNATIVE 3' EXONS.
DR   EMBL; X54648; G804979; -.
DR   EMBL; X54649; G804979; JOINED.
DR   EMBL; X54650; G804979; JOINED.
DR   EMBL; X54651; G804979; JOINED.
DR   EMBL; X54648; G804980; -.
DR   EMBL; X54649; G804980; JOINED.
DR   EMBL; X54650; G804980; JOINED.
DR   EMBL; X54652; G804980; JOINED.
DR   EMBL; X54646; G7981; -.
DR   EMBL; X54647; G7983; ALT_SEQ.
DR   PIR; S03540; S03540.
DR   PIR; S15708; S15708.
DR   FLYBASE; FBgn0001085; fz.
KW   TRANSMEMBRANE; SIGNAL; ALTERNATIVE SPLICING; DEVELOPMENTAL PROTEIN.
FT   SIGNAL         1     26       POTENTIAL.
FT   CHAIN         27    581       FRIZZLED PROTEIN.
FT   TRANSMEM     248    270       POTENTIAL.
FT   TRANSMEM     281    303       POTENTIAL.
FT   TRANSMEM     336    368       POTENTIAL.
FT   TRANSMEM     381    401       POTENTIAL.
```

```
FT   TRANSMEM     422     439       POTENTIAL.
FT   TRANSMEM     472     492       POTENTIAL.
FT   TRANSMEM     530     553       POTENTIAL.
FT   VARSPLIC     406     415       DILSGVCFVG -> MYLWQFHTIN (IN CLASS II).
FT   VARSPLIC     416     581       MISSING (IN CLASS II).
SQ   SEQUENCE   581 AA;  64847 MW;  279EBE85 CRC32;

  Query Match           14.9%;  Score 332;  DB 1;  Length 581;
  Best Local Similarity  37.5%;  Pred. No. 9.66e-53;
  Matches    45;  Conservative   22;  Mismatches  47;  Indels   6;  Gaps   6;

Db      50 HNRCEPI-T-ISICKNIPYNMTIMPNLIGHTKQEEAGLEVHQFAPLVKIGCSDDLQLFLC 107
           :: | || : : :|  | |    :|||:||    |   :     |||:  |  |   |||
Qy      37 RSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 96

Db     108 SLYVPVC-TILERPIPPCRSLCESARV-CEKLMKTYNFNWPENLECSKFPVHGGEDLCVA 165
           ||: |||   |:  | ||:|||   :  |  :| :: | ||: ||| :|| :  :|||::
Qy      97 SLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QD-NDLCIP 154


RESULT    2
ID   NET2_CHICK     STANDARD;      PRT;   581 AA.
AC   Q90923;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   NETRIN-2 PRECURSOR (FRAGMENT).
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WHITE LEGHORN; TISSUE=EMBRYONIC BRAIN;
RX   MEDLINE; 94340732.
RA   SERAFINI T., KENNEDY T.E., GALKO M.J., MIRZAYAN C., JESSELL T.M.,
RA   TESSIER-LAVIGNE M.;
RL   CELL 78:409-424(1994).
CC   -!- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND
CC       PERIPHERAL MOTOR AXONS.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC   -!- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS ONE C345C DOMAIN.
CC   -!- SIMILARITY: TO CAENORHABDITIS ELEGANS UNC-6 PROTEIN.
DR   EMBL; L34550; G529421; -.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 3.
KW   GLYCOPROTEIN; EXTRACELLULAR MATRIX; SIGNAL; LAMININ EGF-LIKE DOMAIN;
KW   REPEAT.
FT   NON_TER       1       1
FT   SIGNAL       <1      15       POTENTIAL.
FT   CHAIN         16     581       NETRIN-2.
FT   DOMAIN        16     261       LAMININ N-TERMINAL (DOMAIN VI).
FT   DOMAIN       262     430       3 X LAMININ EGF-LIKE REPEATS (LAMININ
FT                                  DOMAIN V-LIKE).
FT   DOMAIN       262     317       LAMININ EGF-LIKE 1.
FT   DOMAIN       318     380       LAMININ EGF-LIKE 2.
FT   DOMAIN       381     430       LAMININ EGF-LIKE 3.
FT   DOMAIN       431     581       C345C (DOMAIN C).
FT   SITE         507     509       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID      91     124       BY SIMILARITY.
FT   DISULFID     262     271       BY SIMILARITY.
FT   DISULFID     264     281       BY SIMILARITY.
FT   DISULFID     283     292       BY SIMILARITY.
FT   DISULFID     295     315       BY SIMILARITY.
FT   DISULFID     318     327       BY SIMILARITY.
FT   DISULFID     320     345       BY SIMILARITY.
FT   DISULFID     348     357       BY SIMILARITY.
FT   DISULFID     360     378       BY SIMILARITY.
FT   DISULFID     381     393       BY SIMILARITY.
FT   DISULFID     383     400       BY SIMILARITY.
FT   DISULFID     402     411       BY SIMILARITY.
FT   DISULFID     414     428       BY SIMILARITY.
FT   DISULFID     453     521       BY SIMILARITY.
FT   DISULFID     468     578       BY SIMILARITY.
FT   CARBOHYD      88      88       POTENTIAL.
FT   CARBOHYD     103     103       POTENTIAL.
FT   CARBOHYD     394     394       POTENTIAL.
FT   CARBOHYD     540     540       POTENTIAL.
SQ   SEQUENCE   581 AA;  65106 MW;  5B6D2272 CRC32;

  Query Match           6.6%;  Score 148;  DB 1;  Length 581;
  Best Local Similarity  27.0%;  Pred. No. 8.14e-11;
  Matches    30;  Conservative   28;  Mismatches  46;  Indels   7;  Gaps   6;

Db     431 IPAINPTSLVTSTE-APADCDSYCKPAKGNYKINMKKYCKKDYVVQVNILEMETVANWAK 489
           ||  ::  |: :|| ||  |:: ||  : : :  |   ||:|: : ::: |:  :   :|
Qy     153 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTK 211

Db     490 FTINILS--VYKCRDERVKR-GDNFLWIHLKDLSCKCPKI-QISKKYLVMG 536
           : ::  |  :||      :    : ||:    | | |  : :|:  |||||
Qy     212 IILETKSKTIYKLNGVSERDLKKSVLWLK-DSLQCTCEEMNDINAPYLVMG 261


RESULT    3
ID   NET1_CHICK     STANDARD;      PRT;   606 AA.
AC   Q90922;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   NETRIN-1 PRECURSOR.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WHITE LEGHORN; TISSUE=EMBRYONIC BRAIN;
RX   MEDLINE; 94340732.
RA   SERAFINI T., KENNEDY T.E., GALKO M.J., MIRZAYAN C., JESSELL T.M.,
RA   TESSIER-LAVIGNE M.;
RL   CELL 78:409-424(1994).
CC   -!- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND
CC       PERIPHERAL MOTOR AXONS.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC   -!- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS ONE C345C DOMAIN.
CC   -!- SIMILARITY: TO CAENORHABDITIS ELEGANS UNC-6 PROTEIN.
DR   EMBL; L34549; G529419; -.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 3.
KW   GLYCOPROTEIN; EXTRACELLULAR MATRIX; SIGNAL; LAMININ EGF-LIKE DOMAIN;
KW   REPEAT.
FT   SIGNAL        1      25       POTENTIAL.
FT   CHAIN        26     606       NETRIN-1.
FT   DOMAIN       26     286       LAMININ N-TERMINAL (DOMAIN VI).
FT   DOMAIN      287     455       3 X LAMININ EGF-LIKE REPEATS (LAMININ
FT                                  DOMAIN V-LIKE).
FT   DOMAIN      287     342       LAMININ EGF-LIKE 1.
FT   DOMAIN      343     405       LAMININ EGF-LIKE 2.
FT   DOMAIN      406     455       LAMININ EGF-LIKE 3.
FT   DOMAIN      456     606       C345C (DOMAIN C).
FT   SITE        532     534       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    121     154       BY SIMILARITY.
FT   DISULFID    287     296       BY SIMILARITY.
FT   DISULFID    289     306       BY SIMILARITY.
FT   DISULFID    308     317       BY SIMILARITY.
FT   DISULFID    320     340       BY SIMILARITY.
FT   DISULFID    343     352       BY SIMILARITY.
FT   DISULFID    345     370       BY SIMILARITY.
FT   DISULFID    373     382       BY SIMILARITY.
FT   DISULFID    385     403       BY SIMILARITY.
FT   DISULFID    406     418       BY SIMILARITY.
FT   DISULFID    408     425       BY SIMILARITY.
FT   DISULFID    427     436       BY SIMILARITY.
FT   DISULFID    439     453       BY SIMILARITY.
```

```
FT   DISULFID    478    546       BY SIMILARITY.
FT   DISULFID    493    603       BY SIMILARITY.
FT   CARBOHYD     97     97       POTENTIAL.
FT   CARBOHYD    118    118       POTENTIAL.
FT   CARBOHYD    133    133       POTENTIAL.
FT   CARBOHYD    419    419       POTENTIAL.
SQ   SEQUENCE   606 AA;  68126 MW;  C07CCA06 CRC32;

  Query Match             5.6%;  Score 124;  DB 1;  Length 606;
  Best Local Similarity  26.1%;  Pred. No. 3.06e-06;
  Matches    31;  Conservative   34;  Mismatches  41;  Indels  13;  Gaps  11;

Db    456 IPAAPPPTAASSTEEPAD-CDSYCKASKGKLKINMKKYCKKDYAVQIHILKAEKNADW-W 513
          || |::    ::|||::  |:: ||  :   :  |   ||:|:|: | : |     ::
Qy    153 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDT 210

Db    514 KFTVNIIS--VYK-QG-SNR-LRRGDQTLWVHAKDIACKCPKVKPMKK-YLLLGSTEDS 566
          |: ::  |  :|| | |:| |:::   ||:    : | |  :: ::  ||::|  : :
Qy    211 KIILETKSKTIYKLNGVSERDLKKS--VLWLKD-SLQCTCEEMNDINAPYLVMGQKQGG 266


RESULT    4
ID   YQI8_CAEEL     STANDARD;      PRT;   465 AA.
AC   Q09280;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 53.8 KD PROTEIN C45G9.8 IN CHROMOSOME III.
GN   C45G9.8.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   BENTLEY D., WATERSTON R.;
RL   SUBMITTED (FEB-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U21323; G687873; -.
DR   WORMPEP; C45G9.8; CE01859.
KW   HYPOTHETICAL PROTEIN.
FT   DOMAIN      139    145       POLY-THR.
SQ   SEQUENCE   465 AA;  53756 MW;  37C8121B CRC32;

  Query Match             4.9%;  Score 109;  DB 1;  Length 465;
  Best Local Similarity  33.3%;  Pred. No. 1.32e-03;
  Matches    25;  Conservative   19;  Mismatches  25;  Indels   6;  Gaps   6;

Db    202 QLRPRIEEDQNI-EV-TSRDIDEAD-MERICKK-FAKEKINPITMAEPIDEKTKILLDRV 257
          :| |  ||  :: |   ::: |: | || :||: ||  ||:   ::  |:  |||:|:
Qy    160 HLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFAL-KIKVKEITY-INRDTKIILETK 217

Db    258 TKKPYPLKYNNEKGL 272
          :|  | |:  :|: |
Qy    218 SKTIYKLNGVSERDL 232


RESULT    5
ID   CPD4_RAT       STANDARD;      PRT;   500 AA.
AC   P13108;
DT   01-JAN-1990 (REL. 13, CREATED)
DT   01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   CYTOCHROME P450 IID4 (EC 1.14.14.1) (P450-DB4) (P450-CMF3)
DE   (DEBRISOQUINE 4-HYDROXYLASE).
GN   CYP2D14 OR CYP2D-4.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX   MEDLINE; 90189185.
RA   MATSUNAGA E., UMENO M., GONZALEZ F.J.;
RL   J. MOL. EVOL. 30:155-169(1990).
```

```
RN   [2]
RP   SEQUENCE OF 177-500 FROM N.A.
RX   MEDLINE; 89050091.
RA   ISHIDA N., TAWARAGI Y., INUZUKA C., SUGITA O., KUBOTA I.,
RA   NAKAZATO H., NOGUCHI T., SASSA S.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 156:681-688(1988).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC       OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC       ACIDS, AND XENOBIOTICS.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC   -!- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER
CC       TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,
CC       AND CARCINOGENS.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   EMBL; X52029; G57816; -.
DR   EMBL; M22331; G203830; -.
DR   PIR; D31579; D31579.
DR   PIR; S16873; S16873.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   OXIDOREDUCTASE; MONOOXYGENASE; ELECTRON TRANSPORT; MEMBRANE; HEME;
KW   MICROSOME.
FT   BINDING     446    446       HEME (BY SIMILARITY).
SQ   SEQUENCE   500 AA;  56697 MW;  429F82E7 CRC32;

  Query Match             4.7%;  Score 105;  DB 1;  Length 500;
  Best Local Similarity  24.7%;  Pred. No. 6.13e-03;
  Matches    18;  Conservative   28;  Mismatches  21;  Indels   6;  Gaps   5;

Db    160 ARCLCAAFADHSGFPFS-PNTLLDKAVCNVIASLLFACR-FEYNDPRFIRLL--DLLKDT 215
          ::| | : : :: | |: |:    :: |: |:: |  |: :|| : |:  ||  : :|:
Qy     15 SHC-CLG-SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV 72

Db    216 LEEESGFLPMLLN 228
          ||: :: :|::::
Qy     73 LEQAGAWIPLVMK 85


RESULT    6
ID   CPDI_RAT       STANDARD;      PRT;   500 AA.
AC   Q64680;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   CYTOCHROME P450 IID18 (EC 1.14.14.1) (P450 2D-29/2D-35).
GN   CYP2D18 OR CYP2D-18.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95251650.
RA   KAWASHIMA H., STROBEL H.W.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 209:535-540(1995).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC       OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC       ACIDS, AND XENOBIOTICS.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC   -!- TISSUE SPECIFICITY: BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   EMBL; U48220; G1200518; -.
DR   EMBL; U48219; G1200516; -.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   OXIDOREDUCTASE; MONOOXYGENASE; ELECTRON TRANSPORT; MEMBRANE; HEME;
KW   MICROSOME.
FT   BINDING     446    446       HEME (BY SIMILARITY).
```

```
SQ   SEQUENCE   500 AA;  56683 MW;  F01949B5 CRC32;

  Query Match             4.7%;  Score 105;  DB 1;  Length 500;
  Best Local Similarity  24.7%;  Pred. No. 6.13e-03;
  Matches   18;  Conservative   28;  Mismatches  21;  Indels   6;  Gaps   5;

Db    160 ARCLCAAFADHSGFPFS-PNTLLDKAVCNVIASLLFACR-FEYNDPRFIRLL--DLLKDT 215
          ::| | : : :: | |: |:    :: |: |:: |  |: :|| : |:  ||  : :|:
Qy     15 SHC-CLG-SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV 72

Db    216 LEEESGFLPMLLN 228
          ||: :: :|::::
Qy     73 LEQAGAWIPLVMK 85


RESULT    7
ID   RPC1_GIALA     STANDARD;      PRT;  1741 AA.
AC   P25202;
DT   01-MAY-1992 (REL. 22, CREATED)
DT   01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   DNA-DIRECTED RNA POLYMERASE III LARGEST SUBUNIT (EC 2.7.7.6).
GN   RPOA3.
OS   GIARDIA LAMBLIA (GIARDIA INTESTINALIS).
OC   EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; MASTIGOPHORA; DIPLOMONADIDA;
OC   HEXAMITIDAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SSP. P1;
RX   MEDLINE; 92195823.
RA   LANZENDOERFER M., PALM P., GRAMPP B., PEATTIE D.A., ZILLIG W.;
RL   NUCLEIC ACIDS RES. 20:1145-1145(1992).
CC   -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC       OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC       SUBSTRATES.
CC   -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
CC       RNA(N).
CC   -!- SUBUNIT: RNA POLYMERASE III CONSISTS OF ABOUT 15 DIFFERENT
CC       SUBUNITS. THIS SUBUNIT IS THE LARGEST COMPONENT OF RNA POLYMERASE
CC       III.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE FOUND IN
CC       EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR,
CC       POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE III FOR 5S
CC       AND TRNA GENES.
DR   EMBL; X60325; G9364; -.
KW   DNA-DIRECTED RNA POLYMERASE; TRANSCRIPTION; ZINC; ZINC-FINGER;
KW   NUCLEAR PROTEIN.
FT   ZN_FING      79     92       C2H2-TYPE (POTENTIAL).
SQ   SEQUENCE   1741 AA;  193928 MW;  DCDCD942 CRC32;

  Query Match             4.5%;  Score 101;  DB 1;  Length 1741;
  Best Local Similarity  27.4%;  Pred. No. 2.73e-02;
  Matches   20;  Conservative   25;  Mismatches  23;  Indels   5;  Gaps   5;

Db    168 TICKKSQGFRV-LREMHYEERATKILADVIKHQKDIEKIQELDQR-LRNHVVDPIQALHI 225
          |:||:  :::: ::|: | :| |||: :  |  | | |:  : :| |:: |:   ::|:
Qy    188 TLCKNDFALKIKVKEITYINRDTKIILET-KS-KTIYKLNGVSERDLKKSVLWLKDSLQC 245

Db    226 L-QKVPECDYPYL 237
            : : : : |||
Qy    246 TCEEMNDINAPYL 258


RESULT    8
ID   KS61_MOUSE     STANDARD;      PRT;   724 AA.
AC   P18653;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   RIBOSOMAL PROTEIN S6 KINASE II ALPHA 1 (EC 2.7.1.-) (S6KII-ALPHA)
DE   (P90-RSK 1).
GN   RPS6KA1 OR RSK1.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89384612.
RA   ALCORTA D.A., CREWS C.M., SWEET L.J., BANKSTON L., JONES S.W.,
RA   ERIKSON R.L.;
RL   MOL. CELL. BIOL. 9:3850-3859(1989).
CC   -!- FUNCTION: PHOSPHORYLATES A WIDE RANGE OF SUBSTRATES INCLUDING
CC       RIBOSOMAL PROTEIN S6.
CC   -!- TISSUE SPECIFICITY: INTESTINE, THYMUS, AND LUNG.
CC   -!- SIMILARITY: THE N-TERMINUS SHOWS HIGH SEQUENCE SIMILARITY TO THE
CC       REGIONS OF PROTEIN KINASE C, THE CATALYTIC SUBUNIT OF CAMP-
CC       DEPENDENT PROTEIN KINASE, AND CGMP-DEPENDENT PROTEIN KINASE. THE
CC       REMINDER OF S6 KINASE II SHOWS HIGH SEQUENCE SIMILARITY TO THE
CC       CATALYTIC SUBUNIT OF PHOSPHORYLASE B KINASE.
DR   EMBL; M28489; G556322; -.
DR   PIR; B32571; B32571.
DR   HSSP; P05132; 1CTP.
DR   MGD; MGI:104558; RPS6KA1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 2.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 2.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 2.
KW   TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW   REPEAT; MULTIGENE FAMILY.
FT   DOMAIN       62    310       PROTEIN KINASE 1.
FT   DOMAIN      407    664       PROTEIN KINASE 2.
FT   NP_BIND      68     76       ATP (BY SIMILARITY).
FT   BINDING      94     94       ATP (BY SIMILARITY).
FT   ACT_SITE    187    187       BY SIMILARITY.
FT   NP_BIND     413    421       ATP (BY SIMILARITY).
FT   BINDING     436    436       ATP (BY SIMILARITY).
FT   ACT_SITE    524    524       BY SIMILARITY.
SQ   SEQUENCE   724 AA;  81594 MW;  F961A261 CRC32;

  Query Match             4.4%;  Score 99;  DB 1;  Length 724;
  Best Local Similarity  35.2%;  Pred. No. 5.67e-02;
  Matches   19;  Conservative   13;  Mismatches  18;  Indels   4;  Gaps   4;

Db    489 KI-LRQKFFSEREASFVLHTISKTVEYLHSQGVVHRDLKPSNILYVDESGNPEC 541
          || :::  : :|:: ::| | |||: |    ||  |||| | :| : :|    |
Qy    197 KIKVKEITYINRDTKIILETKSKTI-Y-KLNGVSERDLKKS-VLWLKDSLQCTC 247


RESULT    9
ID   VG07_BPT4      STANDARD;      PRT;  1032 AA.
AC   P19061;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT   01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
DE   BASEPLATE STRUCTURAL PROTEIN GP7.
GN   7.
OS   BACTERIOPHAGE T4.
OC   VIRIDAE; DS-DNA NONENVELOPED VIRUSES; MYOVIRIDAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D;
RX   MEDLINE; 90384864.
RA   EFIMOV V.P., PRILIPOV A.G., MESYANZHINOV V.V.;
RL   NUCLEIC ACIDS RES. 18:5313-5313(1990).
CC   -!- FUNCTION: STRUCTURAL COMPONENT OF THE BASEPLATE.
DR   EMBL; X15907; G15323; -.
DR   PIR; JQ0657; G7BPT4.
KW   STRUCTURAL PROTEIN.
SQ   SEQUENCE   1032 AA;  119214 MW;  39F5B2D1 CRC32;

  Query Match             4.4%;  Score 98;  DB 1;  Length 1032;
  Best Local Similarity  39.0%;  Pred. No. 8.13e-02;
  Matches   16;  Conservative    8;  Mismatches  15;  Indels   2;  Gaps   2;

Db    368 MDSICDKVFALIIGEVETLNANPRTSKII-DSADKGIYVLN 407
          |:::| : ||| |  |  :    | :||| ::  | || ||
```

```
Qy    186 METLCKNDFALKI-KVKEITYINRDTKIILETKSKTIYKLN 225


RESULT   10
ID   ION3_CARAU     STANDARD;      PRT;   520 AA.
AC   P18520;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   INTERMEDIATE FILAMENT PROTEIN ON3.
OS   CARASSIUS AURATUS (GOLDFISH).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC   OSTEICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=GLIAL CELL;
RX   MEDLINE; 90180470.
RA   GIORDANO S., GLASGOW E., TESSER P., SCHECHTER N.;
RL   NEURON 2:1507-1516(1989).
CC   -!- FUNCTION: ONE OF THE NONNEURONAL PREDOMINANT INTERMEDIATE
CC       FILAMENT PROTEINS OF THE VISUAL PATHWAY.
CC   -!- SIMILARITY: SIGNIFICANT HOMOLOGY WITH MAMMALIAN TYPE II KERATIN
CC       K8.
DR   EMBL; M87773; -; NOT_ANNOTATED_CDS.
DR   PIR; JS0291; JS0291.
DR   PROSITE; PS00226; IF; 1.
KW   INTERMEDIATE FILAMENT; COILED COIL; HEPTAD REPEAT PATTERN; VISION.
FT   DOMAIN        1    108       HEAD.
FT   DOMAIN      109    420       ROD.
FT   DOMAIN      421    520       TAIL.
FT   DOMAIN      109    144       COIL 1A.
FT   DOMAIN      145    157       LINKER 1.
FT   DOMAIN      158    253       COIL 1B.
FT   DOMAIN      254    273       LINKER 12.
FT   DOMAIN      274    420       COIL 2.
SQ   SEQUENCE   520 AA;  57789 MW;  3E8CCC03 CRC32;

  Query Match             4.3%;  Score 96;  DB 1;  Length 520;
  Best Local Similarity  32.2%;  Pred. No. 1.66e-01;
  Matches    19;  Conservative   14;  Mismatches  24;  Indels   2;  Gaps   2;

Db    194 KNKYEDEIN-KRTECENDFVLIKKDVDEAYMNKVELEAKLESLSDEINFLRQIFEEEIR 251
          ||| :|: : : | | ||| |  |  : :|:|: :    ||: |  |  |  : | :::
Qy    176 KNKNDDDNDIMETLCKNDFALKIKVKEITYINR-DTKIILETKSKTIYKLNGVSERDLK 233


RESULT   11
ID   UD13_RAT       STANDARD;      PRT;   531 AA.
AC   Q64637;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   UDP-GLUCURONOSYLTRANSFERASE 1-3 PRECURSOR, MICROSOMAL (EC 2.4.1.17)
DE   (UDPGT) (UGT1*3) (UGT1-03) (UGT1.3) (UGT1A3) (B3).
GN   UGT1.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE OF 1-286 FROM N.A.
RC   STRAIN=WISTAR;
RX   MEDLINE; 95332265.
RA   EMI Y., IKUSHIRO S.I., IYANAGI T.;
RL   J. BIOCHEM. 117:392-399(1995).
RN   [2]
RP   SEQUENCE OF 287-531 FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 90274676.
RA   SATO H., KOIWAI O., TANABE K., KASHIWAMATA S.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 169:260-264(1990).
CC   -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC       SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC       ENDOGENOUS COMPOUNDS.
CC   -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC       BETA-D-GLUCORONOSIDE.
CC   -!- SUBCELLULAR LOCATION: MICROSOMAL.
CC   -!- ALTERNATIVE PRODUCTS: THE UGT1 GENE ENCODES FOR MANY DIFFERENT
CC       UDPGT ISOZYMES WHICH HAVE A DIFFERENT N-TERMINAL DOMAIN AND A
CC       COMMON C-TERMINAL DOMAIN OF 245 RESIDUES.
CC   -!- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
DR   EMBL; D38067; G940422; -.
DR   EMBL; M34007; G207579; ALT_TERM.
DR   PROSITE; PS00375; UDPGT; 1.
KW   TRANSFERASE; GLYCOSYLTRANSFERASE; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL;
KW   MULTIGENE FAMILY; MICROSOME; ALTERNATIVE SPLICING.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    531       UDP-GLUCURONOSYLTRANSFERASE 1-3.
FT   TRANSMEM    489    505       POTENTIAL.
FT   CARBOHYD    116    116       POTENTIAL.
FT   CARBOHYD    139    139       POTENTIAL.
FT   CARBOHYD    293    293       POTENTIAL.
FT   CARBOHYD    431    431       POTENTIAL.
SQ   SEQUENCE   531 AA;  60140 MW;  C3B2A846 CRC32;

  Query Match             4.2%;  Score 94;  DB 1;  Length 531;
  Best Local Similarity  35.7%;  Pred. No. 3.35e-01;
  Matches    15;  Conservative   15;  Mismatches   7;  Indels   5;  Gaps   5;

Db    214 PVP-WMYLCH-VNYGSLARLASDLLQREVSVVEILRHASVWL 253
          |:|  : ||| ::| :: ||: :|| :| :: |:| :|: |:
Qy     42 PIPANLQLCHGIEYQNM-RLP-NLLGHE-TMKEVLEQAGAWI 80


RESULT   12
ID   YM81_YEAST     STANDARD;      PRT;  1562 AA.
AC   Q04781; Q04029;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 180.2 KD PROTEIN IN FAA4-HOR7 INTERGENIC REGION.
GN   YMR247C OR YM9408.09C OR YM9920.01C.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE OF 1-956 FROM N.A.
RC   STRAIN=S288C / AB972;
RA   HUNT S., BOWMAN S., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL   SUBMITTED (MAR-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE OF 950-1562 FROM N.A.
RC   STRAIN=S288C / AB972;
RA   GENTLES S., BOWMAN S., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL   SUBMITTED (MAR-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z48639; G732925; -.
DR   EMBL; Z48756; G736313; -.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   1562 AA;  180185 MW;  CEE45672 CRC32;

  Query Match             4.1%;  Score 92;  DB 1;  Length 1562;
  Best Local Similarity  38.8%;  Pred. No. 6.68e-01;
  Matches    19;  Conservative    9;  Mismatches  21;  Indels   0;  Gaps   0;

Db   1397 KNEFDDINSKMDRLTSNDDALTIKLNNITNEVKASYLIDDQKLEISFKL 1445
          ||  || |  |: |  || || ||:::||   : : :| : |    :||
Qy    176 KNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKL 224


RESULT   13
ID   GYRB_ACIS3     STANDARD;      PRT;   216 AA.
AC   Q44273; Q60168;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   DNA GYRASE SUBUNIT B (EC 5.99.1.3) (FRAGMENTS).
GN   GYRB.
OS   ACINETOBACTER SP. (STRAIN ATCC 17924 / CIP 70.12).
```

```
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   NEISSERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97088999.
RA   YAMAMOTO S., HARAYAMA S.;
RL   INT. J. SYST. BACTERIOL. 46:506-511(1996).
CC   -!- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
CC       STRANDED DNA IN AN ATP-DEPENDANT MANNER AND ALSO CATALYZES THE
CC       INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC       DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC   -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
CC       OF DOUBLE-STRANDED DNA.
CC   -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC       BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC       ENZYME FORMS AN A2B2 TETRAMER.
DR   EMBL; D73436; G1322131; -.
DR   EMBL; D73421; G1322101; -.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   TOPOISOMERASE; ISOMERASE; ATP-BINDING; ANTIBIOTIC RESISTANCE.
FT   NON_TER       1      1
FT   NON_CONS    116    117
FT   NON_TER     216    216
SQ   SEQUENCE   216 AA;  23775 MW;  E572EEE0 CRC32;

  Query Match             4.0%;  Score 90;  DB 1;  Length 216;
  Best Local Similarity  26.3%;  Pred. No. 1.31e+00;
  Matches    20;  Conservative   23;  Mismatches  26;  Indels   7;  Gaps   7;

Db      79 LARRLRELSFLNAGVRIVLRDERVALEHIFDLEVGLSEKSAL-DIAGLPGKLA-DCQEKD 136
           |  :::|::::|   :|:| : : :   |: |: |:||:     :  |   |   |:|::
Qy     196 LKIKVKEITYINRDTKIIL-ETK-S-KTIYKLN-GVSERDLKKSVLWLKDSLQCTCEEMN 251

Db     137 PALSELYLVEGDSAGG 152
             ::  ||| |:  ||
Qy     252 -DINAPYLVMGQKQGG 266


RESULT   14
ID   VIRF_SHIDY     STANDARD;      PRT;   262 AA.
AC   Q04248;
DT   01-JUN-1994 (REL. 29, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   VIRULENCE REGULON TRANSCRIPTIONAL ACTIVATOR VIRF.
GN   VIRF.
OS   SHIGELLA DYSENTERIAE, SHIGELLA FLEXNERI, AND SHIGELLA SONNEI.
OG   PLASMID 210 KB INVASION.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.DYSENTERIAE; STRAIN=CG097;
RA   YAO R., REDDY L.V., PALCHAUDHURI S.;
RL   SUBMITTED (MAR-1991) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.FLEXNERI; STRAIN=2A;
RX   MEDLINE; 87032409.
RA   SAKAI T., SASAKAWA C., MAKINO S., YOSHIKAWA M.;
RL   INFECT. IMMUN. 54:395-402(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.SONNEI;
RX   MEDLINE; 89212881.
RA   KATO J.I., ITO K.I., NAKAMURA A., WATANABE H.;
RL   INFECT. IMMUN. 57:1391-1398(1989).
RN   [4]
RP   SIMILARITY TO ARAC FAMILY.
RX   MEDLINE; 92326642.
RA   DORMAN C.J.;
RL   MOL. MICROBIOL. 6:1575-1575(1992).
CC   -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR OF THE VIRB GENE WHICH IS
CC       ITSELF AN ACTIVATOR OF THE IPAABCD VIRULENCE REGULON.
CC   -!- SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL
CC       REGULATORS.
DR   EMBL; X58464; G46954; -.
DR   EMBL; M29172; G152801; -.
DR   EMBL; X16661; G47067; -.
DR   PIR; S14646; S14646.
DR   PIR; A47605; A47605.
DR   PIR; A60105; A60105.
DR   PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
DR   PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW   PLASMID; VIRULENCE; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING.
FT   DNA_BIND    177    196       H-T-H MOTIF (BY SIMILARITY).
SQ   SEQUENCE   262 AA;  30549 MW;  11339B38 CRC32;

  Query Match             4.0%;  Score 89;  DB 1;  Length 262;
  Best Local Similarity  25.5%;  Pred. No. 1.83e+00;
  Matches    13;  Conservative   21;  Mismatches  15;  Indels   2;  Gaps   2;

Db      37 TLTIDEGQIAFIERNIQINVSIKKSDSINPFEIISLDRNLLLSIIRIMEPI 87
              :| |   :|::|:|:  | :  | | :|  :: :| :|:|  |::::::::
Qy     195 ALKIKVKEITYINRDTKIILETK-SKTIYKLNGVS-ERDLKKSVLWLKDSL 243


RESULT   15
ID   KSGA_MYCPN     STANDARD;      PRT;   263 AA.
AC   P75113;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   DIMETHYLADENOSINE TRANSFERASE (EC 2.1.1.-) (S-ADENOSYLMETHIONINE-6-N',
DE   N'-ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RRNA DIMETHYLASE) (HIGH
DE   LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN
DE   DIMETHYLTRANSFERASE).
GN   KSGA.
OS   MYCOPLASMA PNEUMONIAE.
OC   PROKARYOTA; TENERICUTES; MOLLICUTES; MYCOPLASMA; MYCOPLASMATALES;
OC   MYCOPLASMATACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE; 97105885.
RA   HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
RA   HERRMANN R.;
RL   NUCLEIC ACIDS RES. 24:4420-4449(1996).
CC   -!- FUNCTION: SPECIFICALLY DIMETHYLATES TWO ADJACENT ADENOSINES IN THE
CC       LOOP OF A CONSERVED HAIRPIN NEAR THE 3'-END OF 16S RRNA IN THE 30S
CC       PARTICLE. ITS INACTIVATION LEADS TO KASUGAMYCIN RESISTANCE
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: STRONG, TO OTHER BACTERIAL KSGA.
CC   -!- SIMILARITY: BELONGS TO THE RRNA ADENINE N-6-METHYLTRANSFERASE
CC       FAMILY.
DR   EMBL; AE000017; G1673824; -.
DR   PROSITE; PS01131; RRNA_A_DIMETH; 1.
KW   MRNA PROCESSING; TRANSFERASE; METHYLTRANSFERASE;
KW   ANTIBIOTIC RESISTANCE.
SQ   SEQUENCE   263 AA;  29794 MW;  357AA00D CRC32;

  Query Match             4.0%;  Score 90;  DB 1;  Length 263;
  Best Local Similarity  37.0%;  Pred. No. 1.31e+00;
  Matches    20;  Conservative   11;  Mismatches  20;  Indels   3;  Gaps   3;

Db      52 KLQLPYHGIEL-DKRLAEYLLVNEILTEEQLTIGDALKQNLDQYFPDTIPLLCG 104
           :|||  ||||  ::||:: || :| : |     |  :   : |  |||  :||:
Qy      46 NLQL-CHGIEYQNMRLPN-LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCS 97


Search completed: Thu Oct 22 15:51:11 1998
Job time : 26 secs.
```

****

MPSRCH (TM)

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Oct 22 15:53:14 1998; MasPar time 13.37 Seconds
806.104 Million cell updates/sec
Tabular output not generated.

Title: >US-08-848-439-2
Description: (1-295) from US08848439.pep
Perfect Score: 2234
Sequence: 1 MLQGPGSLLLLFLASHCCLG..........WQKGQREFKRISRSIRKLQC 295

Scoring table: PAM 150
Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir56
1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d

Statistics: Mean 44.927; Variance 86.902; scale 0.517

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	386	17.3	641	2	A45054	probable intercellula	1.70e-52
2	355	15.9	694	2	S78444	dfz2 protein - fruit	2.44e-46
3	355	15.9	694	2	S71786	dfz2 protein - fruit	2.44e-46
4	332	14.9	415	2	S15709	hypothetical protein	8.11e-42
5	332	14.9	581	2	S03540	gene frizzled protein	8.11e-42
6	178	8.0	562	2	B56101	collagen alpha 1(XVII	3.36e-13
7	148	6.6	581	2	B54665	netrin-2 precursor -	3.23e-08
8	124	5.6	606	2	A54665	netrin-1 precursor -	1.67e-04
9	122	5.5	793	2	JC5539	Smoothened protein -	3.30e-04
10	105	4.7	500	2	D31579	cytochrome P450 2D4 -	8.40e-02
11	105	4.7	500	2	I52313	cytochrome P450 2D4 h	8.40e-02
12	101	4.5	1735	2	S22812	DNA-directed RNA poly	2.87e-01
13	99	4.4	322	2	H64825	hypothetical protein	5.25e-01
14	98	4.4	391	2	E64474	hypothetical protein	7.08e-01
15	99	4.4	724	2	B32571	ribosomal protein S6	5.25e-01
16	98	4.4	1032	1	G7BPT4	baseplate protein gp7	7.08e-01
17	96	4.3	520	2	JS0291	intermediate filament	1.28e+00
18	96	4.3	735	2	I51901	ribosomal protein S6	1.28e+00
19	97	4.3	15281	2	S41309	cyclosporin synthetas	9.52e-01
20	94	4.2	271	2	E64694	hypothetical protein	2.28e+00
21	93	4.2	735	2	A53300	ribosomal protein S6	3.04e+00
22	92	4.1	203	2	B70254	conserved hypothetica	4.05e+00
23	92	4.1	608	2	B70188	hypothetical protein	4.05e+00
24	92	4.1	801	2	B64679	paralysed flagella pr	4.05e+00
25	92	4.1	1562	2	S53069	probable membrane pro	4.05e+00
26	90	4.0	176	2	S70008	troponin I - Atlantic	7.11e+00
27	89	4.0	262	2	A60105	virF protein - Shigel	9.38e+00
28	89	4.0	262	2	A47605	virF virulence protei	9.38e+00
29	89	4.0	262	2	S14646	virF protein - Shigel	9.38e+00
30	90	4.0	263	2	S73489	S-adenosylmethionine:	7.11e+00
31	89	4.0	508	2	A64571	GMP synthase - Helico	9.38e+00
32	90	4.0	521	2	A53153	glucose transport pro	7.11e+00
33	89	4.0	752	2	A32571	ribosomal protein S6	9.38e+00
34	88	3.9	127	2	S45967	probable membrane pro	1.24e+01
35	88	3.9	182	2	S28310	hypothetical protein	1.24e+01
36	88	3.9	183	2	D69108	phycocyanin alpha phy	1.24e+01
37	88	3.9	370	2	A69430	conserved hypothetica	1.24e+01
38	88	3.9	443	2	JQ1527	alpha-amylase (EC 3.2	1.24e+01
39	88	3.9	445	2	S19990	alpha-amylase (EC 3.2	1.24e+01
40	88	3.9	633	2	C32571	ribosomal protein S6	1.24e+01
41	88	3.9	740	2	I38556	insulin-stimulated pr	1.24e+01
42	88	3.9	849	2	S00030	neurofilament triplet	1.24e+01
43	88	3.9	871	2	I48696	gene Nsk2 protein - m	1.24e+01
44	88	3.9	881	2	I48697	gene Nsk2 protein - m	1.24e+01
45	88	3.9	1102	2	S44772	C29E4.4 protein - Cae	1.24e+01

ALIGNMENTS

```
RESULT      1
ENTRY            A45054     #type complete
TITLE            probable intercellular signal transducer or transmitter Fz-1
                   - rat
ORGANISM         #formal_name Rattus norvegicus #common_name Norway rat
DATE             30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change
                   28-Apr-1995
ACCESSIONS       A45054
REFERENCE        A45054
   #authors      Chan, S.D.; Karpf, D.B.; Fowlkes, M.E.; Hooks, M.; Bradley,
                   M.S.; Vuong, V.; Bambino, T.; Liu, M.Y.; Arnaud, C.D.;
                   Strewler, G.J.; Nissenson, R.A.
   #journal      J. Biol. Chem. (1992) 267:25202-25207
   #title        Two homologs of the Drosophila polarity gene frizzled (fz)
                   are widely expressed in mammalian tissues.
   #cross-references MUID:93094228
   #accession    A45054
      ##status         preliminary; not compared with conceptual translation
      ##molecule_type nucleic acid
      ##residues       1-641 ##label CHA
      ##experimental_source UMR 106 osteosarcoma cell line
      ##note           sequence extracted from NCBI backbone (NCBIP:120154)
SUMMARY          #length 641  #molecular-weight 71054  #checksum 8376

  Query Match            17.3%;  Score 386;  DB 2;  Length 641;
  Best Local Similarity  39.2%;  Pred. No. 1.70e-52;
  Matches    49;  Conservative   27;  Mismatches   43;  Indels    6;  Gaps    6;

Db    108 HGYCQPI-S-IPLCTDIAYNQTIMPNLLGHTNQEDAGLEVHQFYPLVKVQCSAELKFFLC 165
          :: | || : : || | |     :|||||     :    :       |||: || :: | |||
Qy     37 RSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 96

Db    166 SMYAPVC-TVLEQALPPCRSLCERAQG-CEALMNKFGFQWPDTLKCEKFPVHGRGELCVG 223
          |::||||   |:::: ||:|||  :     | ::|: ||| ||| | |::|| :   :||:
Qy     97 SLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QD-NDLCIP 154

Db    224 QNTSD 228
            :||
Qy    155 LASSD 159


RESULT      2
ENTRY            S78444     #type complete
TITLE            dfz2 protein - fruit fly (Drosophila melanogaster)
ORGANISM         #formal_name Drosophila melanogaster
DATE             12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change
                   13-Mar-1998
```

```
ACCESSIONS       S78444
REFERENCE        S78444
   #authors      Bhanot, P.; Wang, Y.; Nathans, J.
   #submission   submitted to the EMBL Data Library, July 1996
   #accession    S78444
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-694 ##label BHA
      ##cross-references EMBL:U65589
SUMMARY          #length 694  #molecular-weight 75437  #checksum 3113

  Query Match            15.9%;  Score 355;  DB 2;  Length 694;
  Best Local Similarity  36.2%;  Pred. No. 2.44e-46;
  Matches    47;  Conservative   32;  Mismatches  45;  Indels   6;  Gaps   6;

Db      51 YGVPAIP-KDPNLRCEEITIPMCRGIGYNMTSFPNEMNHETQDEAGLEVHQFW-PLVEIK 108
           :| | :: | :| :     : :|:|| |    :|| : ||| | || | ||
Qy      28 FGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQ 86

Db     109 CSPDLKFFLCSMYTPICLEDYHKPLPVCRSVCERARSGCAPIMQQYSFEWPERMACEHLP 168
           | || | ||||:::|:||:|    :  |:|:| : :  |||:|  ::| ||: : |:::|
Qy      87 CHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP 146

Db     169 LHGDPDNLCM 178
            : | ::||:
Qy     147 -Q-D-NDLCI 153


RESULT    3
ENTRY            S71786     #type complete
TITLE            dfz2 protein - fruit fly (Drosophila melanogaster)
ORGANISM         #formal_name Drosophila melanogaster
DATE             12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change
                   13-Mar-1998
ACCESSIONS       S71786
REFERENCE        S71786
   #authors      Bhanot, P.; Brink, M.; Harryman Samos, C.; Hsieh, J.C.; Wang,
                   Y.; Macke, J.P.; Andrew, D.; Nathans, J.; Nusse, R.
   #journal      Nature (1996) 382:225-230
   #title        A new member of the frizzled family from Drosophila functions
                   as a wingless receptor.
   #accession    S71786
      ##status        preliminary; nucleic acid sequence not shown
      ##molecule_type DNA
      ##residues      1-694 ##label BHA
      ##cross-references EMBL:U65589
SUMMARY          #length 694  #molecular-weight 75423  #checksum 3095

  Query Match            15.9%;  Score 355;  DB 2;  Length 694;
  Best Local Similarity  36.2%;  Pred. No. 2.44e-46;
  Matches    47;  Conservative   32;  Mismatches  45;  Indels   6;  Gaps   6;

Db      51 YGVPAIP-KDPNLRCEEITIPMCRGIGYNMTSFPNEMNHETQDEAGLEVHQFW-PLVEIK 108
           :| | :: | :| :     : :|:|| |    :|| : ||| | || | ||
Qy      28 FGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQ 86

Db     109 CSPDLKFFLCSMYTPICLEDYHKPLPVCRSVCERARSGCAPIMQQYSFEWPERMACEHLP 168
           | || | ||||:::|:||:|    :  |:|:| : :  |||:|  ::| ||: : |:::|
Qy      87 CHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP 146

Db     169 LHGDPDNLCM 178
            : | ::||:
Qy     147 -Q-D-NDLCI 153


RESULT    4
ENTRY            S15709     #type complete
TITLE            hypothetical protein - fruit fly (Drosophila melanogaster)
ORGANISM         #formal_name Drosophila melanogaster
DATE             21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
                   16-Feb-1997
ACCESSIONS       S15709
REFERENCE        S15708
   #authors      Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
   #journal      Genetics (1990) 126:401-416
   #title        Molecular structure of frizzled, a Drosophila tissue polarity
                   gene.
   #cross-references MUID:91060073
   #accession    S15709
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-415 ##label ADL
      ##cross-references EMBL:X54648
GENETICS
   #gene         FlyBase:fz
      ##cross-references FlyBase:FBgn0001085
   #introns      224/3; 264/3; 329/3; 405/1
SUMMARY          #length 415  #molecular-weight 46117  #checksum 4401

  Query Match            14.9%;  Score 332;  DB 2;  Length 415;
  Best Local Similarity  37.5%;  Pred. No. 8.11e-42;
  Matches    45;  Conservative   22;  Mismatches  47;  Indels   6;  Gaps   6;

Db      50 HNRCEPI-T-ISICKNIPYNMTIMPNLIGHTKQEEAGLEVHQFAPLVKIGCSDDLQLFLC 107
           :: | || : : :|  | |    :|||:||    |   :    |||:  |  |   |||
Qy      37 RSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 96

Db     108 SLYVPVC-TILERPIPPCRSLCESARV-CEKLMKTYNFNWPENLECSKFPVHGGEDLCVA 165
           ||: |||   |:  | ||:|||   :  |  :| :: | ||: ||| :|| :  :|||::
Qy      97 SLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QD-NDLCIP 154


RESULT    5
ENTRY            S03540     #type complete
TITLE            gene frizzled protein precursor - fruit fly (Drosophila
                   melanogaster)
ORGANISM         #formal_name Drosophila melanogaster
DATE             31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
                   23-Aug-1997
ACCESSIONS       S03540; S15708
REFERENCE        S03540
   #authors      Vinson, C.R.; Conover, S.; Adler, P.N.
   #journal      Nature (1989) 338:263-264
   #title        A Drosophila tissue polarity locus encodes a protein
                   containing seven potential transmembrane domains.
   #cross-references MUID:89159415
   #accession    S03540
      ##status        not compared with conceptual translation
      ##molecule_type mRNA
      ##residues      1-581 ##label VIN
REFERENCE        S15708
   #authors      Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
   #journal      Genetics (1990) 126:401-416
   #title        Molecular structure of frizzled, a Drosophila tissue polarity
                   gene.
   #cross-references MUID:91060073
   #accession    S15708
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-581 ##label ADL
      ##cross-references EMBL:X54648
GENETICS
   #gene         FlyBase:fz
      ##cross-references FlyBase:FBgn0001085
   #introns      224/3; 264/3; 329/3; 405/1
KEYWORDS         alternative splicing; transmembrane protein
FEATURE
   1-26               #domain signal sequence #status predicted #label SIG\
   27-581             #product gene frizzled protein #status predicted #label
                        MAT
SUMMARY          #length 581  #molecular-weight 64846  #checksum 358

  Query Match            14.9%;  Score 332;  DB 2;  Length 581;
  Best Local Similarity  37.5%;  Pred. No. 8.11e-42;
  Matches    45;  Conservative   22;  Mismatches  47;  Indels   6;  Gaps   6;
```

```
Db      50 HNRCEPI-T-ISICKNIPYNMTIMPNLIGHTKQEEAGLEVHQFAPLVKIGCSDDLQLFLC 107
           :: | || : : :|  | |    :|||:||    |   :     |||:  |  |   |||
Qy      37 RSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 96

Db     108 SLYVPVC-TILERPIPPCRSLCESARV-CEKLMKTYNFNWPENLECSKFPVHGGEDLCVA 165
           ||: |||   |:  | ||:|||   :  |  :| :: | ||: ||| :|| :  :|||::
Qy      97 SLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QD-NDLCIP 154


RESULT    6
ENTRY             B56101     #type fragment
TITLE             collagen alpha 1(XVIII) chain precursor long form - mouse
                    (fragment)
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change
                    10-Sep-1997
ACCESSIONS        B56101
REFERENCE         A56101
   #authors       Rehn, M.; Pihlajaniemi, T.
   #journal       J. Biol. Chem. (1995) 270:4705-4711
   #title         Identification of three N-terminal ends of type XVIII
                    collagen chains and tissue-specific differences in the
                    expression of the corresponding transcripts. The longest
                    form contains a novel motif homologous to rat and
                    Drosophila frizzled proteins.
   #accession     B56101
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-562 ##label REH
      ##cross-references GB:U11637; NID:g618429; PID:g618430
GENETICS
   #gene          COL18A1
SUMMARY           #length 562  #checksum 115

  Query Match             8.0%;  Score 178;  DB 2;  Length 562;
  Best Local Similarity  34.1%;  Pred. No. 3.36e-13;
  Matches    30;  Conservative   13;  Mismatches  45;  Indels   0;  Gaps   0;

Db     368 SRCLPLPPTLTLCSRLGIGHFWLPNHLHHTDSVEVEATVQAWGRFLHTNCHPFLAWFFCL 427
           | | |:|: | ||  :   : :|||  | |    ||     ||  ::   |||    |:|
Qy      38 SNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCS 97

Db     428 LLAPSCGPGPPPPLPPCRQFCEALEDEC 455
           |:|| |       : ||: :|   : | |
Qy      98 LFAPVCLDDLDETIQPCHSLCVQVKDRC 125


RESULT    7
ENTRY             B54665     #type fragment
TITLE             netrin-2 precursor - chicken (fragment)
ORGANISM          #formal_name Gallus gallus #common_name chicken
DATE              06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
                    10-Sep-1997
ACCESSIONS        B54665
REFERENCE         A54665
   #authors       Serafini, T.; Kennedy, T.E.; Galko, M.J.; Mirzayan, C.;
                    Jessell, T.M.; Tessier-Lavigne, M.
   #journal       Cell (1994) 78:409-424
   #title         The netrins define a family of axon outgrowth-promoting
                    proteins homologous to Caenorhabditis elegans UNC-6.
   #accession     B54665
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-581 ##label SER
      ##cross-references GB:L34550; NID:g529420; PID:g529421
SUMMARY           #length 581  #checksum 7731

  Query Match             6.6%;  Score 148;  DB 2;  Length 581;
  Best Local Similarity  27.0%;  Pred. No. 3.23e-08;
  Matches    30;  Conservative   28;  Mismatches  46;  Indels   7;  Gaps   6;

Db     431 IPAINPTSLVTSTE-APADCDSYCKPAKGNYKINMKKYCKKDYVVQVNILEMETVANWAK 489
           ||  ::  |: :|| ||  |:: ||  : : :  |   ||:|: : ::: |:  :   :|
```

```
Qy     153 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTK 211

Db     490 FTINILS--VYKCRDERVKR-GDNFLWIHLKDLSCKCPKI-QISKKYLVMG 536
           : ::  |  :||      :    : ||:    | | |  : :|:  |||||
Qy     212 IILETKSKTIYKLNGVSERDLKKSVLWLK-DSLQCTCEEMNDINAPYLVMG 261


RESULT    8
ENTRY             A54665     #type complete
TITLE             netrin-1 precursor - chicken
ORGANISM          #formal_name Gallus gallus #common_name chicken
DATE              06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
                    10-Sep-1997
ACCESSIONS        A54665
REFERENCE         A54665
   #authors       Serafini, T.; Kennedy, T.E.; Galko, M.J.; Mirzayan, C.;
                    Jessell, T.M.; Tessier-Lavigne, M.
   #journal       Cell (1994) 78:409-424
   #title         The netrins define a family of axon outgrowth-promoting
                    proteins homologous to Caenorhabditis elegans UNC-6.
   #accession     A54665
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-606 ##label SER
      ##cross-references GB:L34549; NID:g529418; PID:g529419
SUMMARY           #length 606  #molecular-weight 68126  #checksum 5400

  Query Match             5.6%;  Score 124;  DB 2;  Length 606;
  Best Local Similarity  26.1%;  Pred. No. 1.67e-04;
  Matches    31;  Conservative   34;  Mismatches  41;  Indels  13;  Gaps  11;

Db     456 IPAAPPPTAASSTEEPAD-CDSYCKASKGKLKINMKKYCKKDYAVQIHILKAEKNADW-W 513
           || |::    ::|||::  |:: ||  :   :  |   ||:|:|: | : |     ::
Qy     153 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDT 210

Db     514 KFTVNIIS--VYK-QG-SNR-LRRGDQTLWVHAKDIACKCPKVKPMKK-YLLLGSTEDS 566
           |: ::  |  :||  | |:| |:::   ||:    : | |  :: ::  ||::|  : :
Qy     211 KIILETKSKTIYKLNGVSERDLKKS--VLWLKD-SLQCTCEEMNDINAPYLVMGQKQGG 266


RESULT    9
ENTRY             JC5539     #type complete
TITLE             Smoothened protein - mouse
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change
                    05-Sep-1997
ACCESSIONS        JC5539; PC4476
REFERENCE         JC5539
   #authors       Akiyama, H.; Shigeno, C.; Hiraki, Y.; Shukunami, C.; Kohno,
                    H.; Akagi, M.; Konishi, J.; Nakamura, T.
   #journal       Biochem. Biophys. Res. Commun. (1997) 235:142-147
   #title         Cloning of a mouse Smoothened cDNA and expression patterns of
                    hedgehog signalling molecules during chondrogenesis and
                    cartilage differentiation in clonal mouse EC cells, ATDC5.
   #accession     JC5539
      ##molecule_type mRNA
      ##residues      1-793 ##label AKI
   #accession     PC4476
      ##molecule_type protein
      ##residues      528-533;539-545;600-605 ##label AK2
      ##experimental_source ADTC5 cell
COMMENT     This protein is used in the conserved targets in Hedgehog
              signalling pathway, together with Patched and Gli. These protein
              are responsible for the skeletal abnormalities in Gorlin and
              Greig syndromes.
FEATURE
   1-32               #domain signal sequence #status predicted #label SIG
SUMMARY           #length 793  #molecular-weight 87299  #checksum 8609

  Query Match             5.5%;  Score 122;  DB 2;  Length 793;
  Best Local Similarity  26.7%;  Pred. No. 3.30e-04;
  Matches    31;  Conservative   22;  Mismatches  54;  Indels   9;  Gaps   7;
```

```
Db      70 RAAHCEPLRYNVCLGSALPYGATT-TLLAGDSDSQEEAHGKLVLWSGLRNAPRCWAVIQP 128
           : ::| |:  |: |  :: |       | |  ::  |       |  |    :| :
Qy      36 KRSNCKPIPANLQLCHGIEYQNMRLPNLLGH-ETMKEVLEQAGAWIPL-VMKQCHPDTKK 93

Db     129 LLCAVYMPKC-EN-DRVELPSRTLCQATRGPCAIVERERG--WPDFLRCTPDHFPE 180
           :||::: | | :: |    | ::||   :  || |    |  ||| | |  |:||:
Qy      94 FLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLEC--DRFPQ 147


RESULT   10
ENTRY            D31579     #type complete
TITLE            cytochrome P450 2D4 - rat
ALTERNATE_NAMES  cytochrome P450CMF3
ORGANISM         #formal_name Rattus norvegicus #common_name Norway rat
DATE             28-Feb-1990 #sequence_revision 17-May-1996 #text_change
                   23-Jan-1998
ACCESSIONS       S16873; D31579
REFERENCE        S16871
   #authors      Matsunaga, E.; Umeno, M.; Gonzalez, F.J.
   #journal      J. Mol. Evol. (1990) 30:155-169
   #title        The rat P450 IID subfamily: complete sequences of four
                   closely linked genes and evidence that gene conversions
                   maintained sequence homogeneity at the heme-binding region
                   of the cytochrome P450 active site.
   #cross-references MUID:90189185
   #accession    S16873
      ##molecule_type DNA
      ##residues      1-500 ##label MAT
      ##cross-references EMBL:X52029; NID:g57815; PID:g57816
REFERENCE        A90151
   #authors      Ishida, N.; Tawaragi, Y.; Inuzuka, C.; Sugita, O.; Kubota,
                   I.; Nakazato, H.; Noguchi, T.; Sassa, S.
   #journal      Biochem. Biophys. Res. Commun. (1988) 156:681-688
   #title        Four species of cDNAs for cytochrome P450 isozymes
                   immunorelated to P450c-M/F encode for members of P450IID
                   subfamily, increasing the number of members within the
                   subfamily.
   #cross-references MUID:89050091
   #accession    D31579
      ##molecule_type mRNA
      ##residues      177-500 ##label ISH
      ##cross-references EMBL:M22331; NID:g203829; PID:g203830
GENETICS
   #gene         CYP2D4
   #introns      63/3; 121/1; 172/1; 225/3; 284/3; 332/1; 394/3; 442/1
CLASSIFICATION   #superfamily cytochrome P450
KEYWORDS         chromoprotein; electron transfer; endoplasmic reticulum;
                   heme; iron; monooxygenase; oxidoreductase; transmembrane
                   protein
FEATURE
   446               #binding_site heme iron (Cys) (axial ligand) #status
                       predicted
SUMMARY          #length 500  #molecular-weight 56697  #checksum 7863

  Query Match             4.7%;  Score 105;  DB 2;  Length 500;
  Best Local Similarity  24.7%;  Pred. No. 8.40e-02;
  Matches    18;  Conservative   28;  Mismatches  21;  Indels   6;  Gaps   5;

Db     160 ARCLCAAFADHSGFPFS-PNTLLDKAVCNVIASLLFACR-FEYNDPRFIRLL--DLLKDT 215
           ::| | : : :: | |: |:    :: |: |:: |  |: :|| : |:  ||  : :|:
Qy      15 SHC-CLG-SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV 72

Db     216 LEEESGFLPMLLN 228
           ||: :: :|::::
Qy      73 LEQAGAWIPLVMK 85


RESULT   11
ENTRY            I52313     #type complete
TITLE            cytochrome P450 2D4 homolog, brain (clone 2d-29) - rat
ORGANISM         #formal_name Rattus norvegicus #common_name Norway rat
DATE             27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change
                   23-Jan-1998
ACCESSIONS       I52313
REFERENCE        I52313
   #authors      Kawashima, H.; Strobel, H.W.
   #journal      Biochem. Biophys. Res. Commun. (1995) 209:535-540
   #title        cDNA cloning of a novel rat brain cytochrome P450 belonging
                   to the CYP2D subfamily.
   #cross-references MUID:95251650
   #accession    I52313
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues      1-500 ##label RES
      ##cross-references GB:S77859; NID:g998524; PID:g998525
      ##experimental_source brain, strain Sprague-Dawley
CLASSIFICATION   #superfamily cytochrome P450
KEYWORDS         heme; transmembrane protein
FEATURE
   446               #binding_site heme iron (Cys) (axial ligand) #status
                       predicted
SUMMARY          #length 500  #molecular-weight 56683  #checksum 8540

  Query Match             4.7%;  Score 105;  DB 2;  Length 500;
  Best Local Similarity  24.7%;  Pred. No. 8.40e-02;
  Matches    18;  Conservative   28;  Mismatches  21;  Indels   6;  Gaps   5;

Db     160 ARCLCAAFADHSGFPFS-PNTLLDKAVCNVIASLLFACR-FEYNDPRFIRLL--DLLKDT 215
           ::| | : : :: | |: |:    :: |: |:: |  |: :|| : |:  ||  : :|:
Qy      15 SHC-CLG-SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV 72

Db     216 LEEESGFLPMLLN 228
           ||: :: :|::::
Qy      73 LEQAGAWIPLVMK 85


RESULT   12
ENTRY            S22812     #type complete
TITLE            DNA-directed RNA polymerase (EC 2.7.7.6) III largest chain -
                   Giardia lamblia
ORGANISM         #formal_name Giardia lamblia
DATE             12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
                   12-Sep-1997
ACCESSIONS       S22812
REFERENCE        S22811
   #authors      Lanzendoerfer, M.; Palm, P.; Grampp, B.; Peattie, D.A.;
                   Zillig, W.
   #journal      Nucleic Acids Res. (1992) 20:1145
   #title        Nucleotide sequence of the gene encoding the largest subunit
                   of the DNA-dependent RNA polymerase III of Giardia lamblia.
   #cross-references MUID:92195823
   #accession    S22812
      ##status        nucleic acid sequence not shown; translation not shown
      ##molecule_type DNA
      ##residues      1-1735 ##label LAN
      ##cross-references EMBL:X60325
      ##note          the nucleotide sequence was submitted to the EMBL Data
                       Library, September 1991
GENETICS
   #gene         rpoA3
CLASSIFICATION   #superfamily Giardia DNA-directed RNA polymerase III largest
                   chain
KEYWORDS         DNA binding; nucleotidyltransferase; nucleus; transcription;
                   zinc finger
SUMMARY          #length 1735  #molecular-weight 193191  #checksum 769

  Query Match             4.5%;  Score 101;  DB 2;  Length 1735;
  Best Local Similarity  27.4%;  Pred. No. 2.87e-01;
  Matches    20;  Conservative   25;  Mismatches  23;  Indels   5;  Gaps   5;

Db     168 TICKKSQGFRV-LREMHYEERATKILADVIKHQKDIEKIQELDQR-LRNHVVDPIQALHI 225
           |:||:  :::: ::|: | :| |||: :  |  | | |:  : :| |:: |:   ::|:
Qy     188 TLCKNDFALKIKVKEITYINRDTKIILET-KS-KTIYKLNGVSERDLKKSVLWLKDSLQC 245

Db     226 L-QKVPECDYPYL 237
           : : : : |||
```

```
Qy    246 TCEEMNDINAPYL 258


RESULT   13
ENTRY            H64825     #type complete
TITLE            hypothetical protein b0872 - Escherichia coli (strain K-12)
ORGANISM         #formal_name Escherichia coli
DATE             12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change
                   14-Nov-1997
ACCESSIONS       H64825
REFERENCE        A64720
   #authors      Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
                   Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
                   Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.;
                   Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao,
                   Y.
   #journal      Science (1997) 277:1453-1462
   #title        The complete genome sequence of Escherichia coli K-12.
   #cross-references MUID:97426617
   #accession    H64825
      ##status        preliminary; nucleic acid sequence not shown;
                       translation not shown
      ##molecule_type DNA
      ##residues      1-322 ##label BLAT
      ##cross-references GB:AE000189; GB:U00096; NID:g1787097; PID:g1787098;
                       UWGP:b0872
      ##experimental_source strain K-12, substrain MG1655
SUMMARY          #length 322  #molecular-weight 35740  #checksum 5190

  Query Match             4.4%;  Score 99;  DB 2;  Length 322;
  Best Local Similarity  27.5%;  Pred. No. 5.25e-01;
  Matches    19;  Conservative   22;  Mismatches  21;  Indels   7;  Gaps   7;

Db     87 RDVKRGDYLWLSDAMGEFTCDDKAE-DK-FLLLAA-GCGVTPIMSMRRWLAKNRPQADVR 143
          ||:|::  ||| |:: : ||::: : :  :|:::    |   | |::||  | : :   |
Qy    230 RDLKKSV-LWLKDSL-QCTCEEMNDINAPYLVMGQKQGGELVITSVKRW-QKGQREFK-R 285

Db    144 VIYNVRTPQ 152
          :  ::|  |
Qy    286 ISRSIRKLQ 294


RESULT   14
ENTRY            E64474     #type complete
TITLE            hypothetical protein MJ1398 - Methanococcus jannaschii
ORGANISM         #formal_name Methanococcus jannaschii
DATE             13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
                   10-Oct-1997
ACCESSIONS       E64474
REFERENCE        A64300
   #authors      Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,
                   R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton,
                   R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.;
                   Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;
                   Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
                   Scott, J.L.; Geoghagen, N.S.M.; Weidman, J.F.; Fuhrmann,
                   J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson,
                   J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,
                   K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk,
                   H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
   #journal      Science (1996) 273:1058-1073
   #title        Complete genome sequence of the methanogenic archaeon,
                   Methanococcus jannaschii.
   #cross-references MUID:96337999
   #accession    E64474
      ##status        preliminary; nucleic acid sequence not shown;
                       translation not shown
      ##molecule_type DNA
      ##residues      1-391 ##label BUL
      ##cross-references GB:U67580; GB:L77117; NID:g1592044; PID:g1592046;
                       TIGR:MJ1398; PID:g1511417
GENETICS
   #map_position FOR1358185-1359360

SUMMARY          #length 391  #molecular-weight 45529  #checksum 6984

  Query Match             4.4%;  Score 98;  DB 2;  Length 391;
  Best Local Similarity  30.0%;  Pred. No. 7.08e-01;
  Matches    15;  Conservative   19;  Mismatches  14;  Indels   2;  Gaps   2;

Db     84 LRFKIENKSYNLIITKKNVNTNKMAIIFTSINGSREVNGNKTIIWIKDPL 133
          |::|: : :|    ||  ::|: ::| :  :||  | : :|:::|:||:|
Qy    196 LKIKVKEITYINRDTKIILETKSKTI-YK-LNGVSERDLKKSVLWLKDSL 243


RESULT   15
ENTRY            B32571     #type complete
TITLE            ribosomal protein S6 kinase II (EC 2.7.-.-) alpha chain
                   homolog (clone Mu6A) - mouse
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
                   26-Feb-1998
ACCESSIONS       B32571
REFERENCE        A93113
   #authors      Alcorta, D.A.; Crews, C.M.; Sweet, L.J.; Bankston, L.; Jones,
                   S.W.; Erikson, R.L.
   #journal      Mol. Cell. Biol. (1989) 9:3850-3859
   #title        Sequence and expression of chicken and mouse rsk: homologs of
                   Xenopus laevis ribosomal S6 kinase.
   #cross-references MUID:89384612
   #accession    B32571
      ##molecule_type mRNA
      ##residues      1-724 ##label ALC
      ##cross-references GB:M23489
CLASSIFICATION   #superfamily ribosomal protein S6 kinase II; protein kinase
                   homology
KEYWORDS         ATP; phosphotransferase; serine/threonine-specific protein
                   kinase
FEATURE
   60-310             #domain protein kinase homology #label KIN1\
   68-76              #region protein kinase ATP-binding motif\
   405-664            #domain protein kinase homology #label KIN2
SUMMARY          #length 724  #molecular-weight 81594  #checksum 4206

  Query Match             4.4%;  Score 99;  DB 2;  Length 724;
  Best Local Similarity  35.2%;  Pred. No. 5.25e-01;
  Matches    19;  Conservative   13;  Mismatches  18;  Indels   4;  Gaps   4;

Db    489 KI-LRQKFFSEREASFVLHTISKTVEYLHSQGVVHRDLKPSNILYVDESGNPEC 541
          || :::  : :|:: ::| | |||: |    ||  |||| | :| : :|    |
Qy    197 KIKVKEITYINRDTKIILETKSKTI-Y-KLNGVSERDLKKS-VLWLKDSLQCTC 247


Search completed: Thu Oct 22 15:54:26 1998
Job time : 72 secs.
```

****

MPSRCH (TM)

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Oct 22 15:57:27 1998; MasPar time 17.59 Seconds
778.674 Million cell updates/sec

Tabular output not generated.

Title: >US-08-848-439-3
Description: (1-275) from US08848439.pep
Perfect Score: 2073
Sequence: 1 SARGLFLFGQPDFSYKRSNC.....WQKGQREFKRISRSIRKLQC 275

Scoring table: PAM 150
Gap 11

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembl6
1:sp_archea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 44.690; Variance 72.884; scale 0.613

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	2067	99.7	295	11	O08862	SECRETED FRIZZLED RELA	0.00e+00
2	2065	99.6	295	11	O35297	SECRETED APOPTOSIS REL	0.00e+00
3	2062	99.5	295	11	P97299	STROMAL CELL DERIVED F	0.00e+00
4	1377	66.4	206	4	O14778	SECRETED APOPTOSIS REL	0.00e+00
5	913	44.0	314	4	O14779	SECRETED APOPTOSIS REL	3.09e-193
6	911	43.9	313	4	O00546	SECRETED FRIZZLED-RELA	9.94e-193
7	902	43.5	308	6	O19116	FRZA PRECURSOR.	1.91e-190
8	901	43.5	317	4	O14780	SECRETED APOPTOSIS REL	3.42e-190
9	890	42.9	314	11	O08861	SECRETED FRIZZLED RELA	2.10e-187
10	729	35.2	307	13	O42397	CRESCENT.	8.17e-147
11	582	28.1	281	13	O73821	SECRETED XWNT8 INHIBIT	3.03e-110
12	398	19.2	568	5	O16147	MOM-5.	1.38e-65
13	397	19.2	592	13	O57328	7-TRANSMEMBRANE PROTEI	2.38e-65
14	396	19.1	572	11	Q61090	FRIZZLED HOMOLOG 7 (TR	4.11e-65
15	394	19.0	591	4	O00144	FRIZZLED HOMOLOG.	1.23e-64
16	394	19.0	685	11	Q61091	FRIZZLED HOMOLOG 8 (TR	1.23e-64
17	391	18.9	567	13	O57329	7-TRANSMEMBRANE PROTEI	6.31e-64
18	392	18.9	626	11	O70421	FRIZZLED-1.	3.65e-64
19	386	18.6	565	4	Q14332	FRIZZLED GENE PRODUCT.	9.65e-63
20	386	18.6	570	11	Q08464	FRIZZLED PROTEIN HOMOL	9.65e-63
21	386	18.6	641	11	Q08463	FRIZZLED PROTEIN HOMOL	9.65e-63
22	374	18.0	585	4	Q13467	TRANSMEMBRANE RECEPTOR	6.60e-60
23	369	17.8	537	11	Q61088	FRIZZLED HOMOLOG 4 (TR	9.94e-59
24	357	17.2	525	5	Q10662	FRIZZLED PROTEIN HOMOL	6.54e-56
25	355	17.1	694	5	Q94916	DFZ2.	1.92e-55
26	347	16.7	586	5	Q24760	TISSUE POLARITY PROTEI	1.43e-53
27	337	16.3	558	5	Q94132	TRANSMEMBRANE RECEPTOR	3.05e-51
28	317	15.3	666	11	Q61086	FRIZZLED-3.	1.30e-46
29	309	14.9	664	13	O42579	FRIZZLED 3 PROTEIN.	8.94e-45
30	295	14.2	605	5	O16190	F27E11.3 PROTEIN.	1.42e-41
31	293	14.1	325	6	Q95117	FRZB PRECURSOR.	4.04e-41
32	291	14.0	325	4	O00181	FRITZ.	1.15e-40
33	289	13.9	325	4	Q92765	FRZB PRECURSOR.	3.27e-40
34	289	13.9	325	4	Q99686	FREZZLED.	3.27e-40
35	284	13.7	709	11	Q61089	FRIZZLED HOMOLOG 6 (TR	4.44e-39
36	281	13.6	348	11	O35222	FRIZZLED RELATED PROTE	2.12e-38
37	275	13.3	323	11	P97401	SECRETED FRIZZLED-RELA	4.76e-37
38	274	13.2	346	4	O14877	FRPHE.	7.98e-37
39	267	12.9	319	13	P79993	FRZB PRECURSOR.	2.96e-35
40	268	12.9	706	4	O60353	FRIZZLED-6.	1.77e-35
41	265	12.8	318	13	P79936	FREZZLED.	8.29e-35
42	240	11.6	261	11	O08570	FRITZ (FRAGMENT).	2.81e-29
43	195	9.4	641	4	O00520	CARBOXYPEPTIDASE Z PRE	1.18e-19
44	178	8.6	505	11	Q60569	PROCOLLAGEN, TYPE XVII	3.68e-16
45	178	8.6	562	11	Q60672	PROCOLLAGEN, TYPE XVII	3.68e-16

ALIGNMENTS

RESULT 1
ID O08862 PRELIMINARY; PRT; 295 AA.
AC O08862;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE SECRETED FRIZZLED RELATED PROTEIN SFRP-2.
GN SFRP2.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97250455.
RA RATTNER A., HSIEH J.C., SMALLWOOD P.M., GILBERT D.J., COPELAND N.G.,
RA JENKINS N.A., NATHANS J.;
RL PROC. NATL. ACAD. SCI. U.S.A. 94:2859-2863(1997).
DR EMBL; U88567; G1946343; -.
SQ SEQUENCE 295 AA; 33469 MW; 236B0DDA CRC32;

Query Match 99.7%; Score 2067; DB 11; Length 295;
Best Local Similarity 99.3%; Pred. No. 0.00e+00;
Matches 273; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Db     21 SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 60

Db     81 PLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDML 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 PLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDML 120

Db    141 ECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKTKNEDDNDIMETLCKNDFALKIKV 200
           |||||||||||||||||||||||||||||||| ||:||||||||||||||||||||||||
Qy    121 ECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKV 180

Db    201 KEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVM 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 KEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVM 240

Db    261 GQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 295
           ||||||||||||||||||||||||||||||||||
Qy    241 GQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 275
```

```
RESULT    2
ID   O35297      PRELIMINARY;      PRT;   295 AA.
AC   O35297;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED APOPTOSIS RELATED PROTEIN 1.
GN   SARP1.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M.,
RA   FITZPATRIC P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF017989; G2415421; -.
SQ   SEQUENCE   295 AA;  33483 MW;  0E4C8A8E CRC32;

  Query Match           99.6%;  Score 2065;  DB 11;  Length 295;
  Best Local Similarity  98.9%;  Pred. No. 0.00e+00;
  Matches   272;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db      21 SARGLFLFGQPDFSYKRTNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 80
           |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Qy       1 SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 60

Db      81 PLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDML 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      61 PLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDML 120

Db     141 ECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKTKNEDDNDIMETLCKNDFALKIKV 200
           |||||||||||||||||||||||||||||||| ||:||||||||||||||||||||
Qy     121 ECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKV 180

Db     201 KEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVM 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     181 KEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVM 240

Db     261 GQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 295
           |||||||||||||||||||||||||||||||||||
Qy     241 GQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 275


RESULT    3
ID   P97299      PRELIMINARY;      PRT;   295 AA.
AC   P97299;
DT   01-MAY-1997 (TREMBLREL. 03, CREATED)
DT   01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   STROMAL CELL DERIVED FACTOR 5 (SDF5).
GN   SDF5.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97092876.
RA   SHIROZU M., TADA H., TASHIRO K., NAKAMURA T., LOPEZ N.D., NAZAREA M.,
RA   HAMADA T., SATO T., NAKANO T., HONJO T.;
RL   GENOMICS 37:273-280(1996).
DR   EMBL; D50462; G1747302; -.
DR   MGD; MGI:108078; SDF5.
SQ   SEQUENCE   295 AA;  33501 MW;  329888C8 CRC32;

  Query Match           99.5%;  Score 2062;  DB 11;  Length 295;
  Best Local Similarity  98.9%;  Pred. No. 0.00e+00;
  Matches   272;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db      21 SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1 SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 60

Db      81 PLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCMQVKDRCAPVMSAFGFPWPDML 140
           ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Qy      61 PLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDML 120

Db     141 ECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKTKNEDDNDIMETLCKNDFALKIKV 200
           |||||||||||||||||||||||||||||||| ||:||||||||||||||||||||
Qy     121 ECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKV 180

Db     201 KEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVM 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     181 KEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVM 240

Db     261 GQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 295
           |||||||||||||||||||||||||||||||||||
Qy     241 GQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 275


RESULT    4
ID   O14778      PRELIMINARY;      PRT;   206 AA.
AC   O14778;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED APOPTOSIS RELATED PROTEIN 1 (FRAGMENT).
GN   SARP1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M.,
RA   FITZPATRIC P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF017986; G2415415; -.
FT   NON_TER     206     206
SQ   SEQUENCE   206 AA;  23215 MW;  F465B68B CRC32;

  Query Match           66.4%;  Score 1377;  DB 4;  Length 206;
  Best Local Similarity  98.4%;  Pred. No. 0.00e+00;
  Matches   185;  Conservative    0;  Mismatches   1;  Indels   2;  Gaps   1;

Db      21 SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1 SARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 60

Db      81 PLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSRCVQVKDRCAPVMSAF--PWPDML 138
           |||||||||||||||||||||||||||||||||| |||||||||||||||  ||||||
Qy      61 PLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDML 120

Db     139 ECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKV 198
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     121 ECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKV 180

Db     199 KEITYINR 206
           ||||||||
Qy     181 KEITYINR 188


RESULT    5
ID   O14779      PRELIMINARY;      PRT;   314 AA.
AC   O14779;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED APOPTOSIS RELATED PROTEIN 2.
GN   SARP2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
```

```
RA   MELKONYAN H., PROCHAZKA V., CHANG W.C., UMANSKY S.R.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF017987; G2415417; -.
SQ   SEQUENCE   314 AA;  35411 MW;  61E261B5 CRC32;

  Query Match           44.0%;  Score 913;  DB 4;  Length 314;
  Best Local Similarity 42.6%;  Pred. No. 3.09e-193;
  Matches  112;  Conservative   67;  Mismatches  77;  Indels   7;  Gaps   5;

Db      51 FYTKPPQCVDIPADLRLCHNVGYKKMVLPNLLEHETMAEVKQQASSWVPLLNKNCHAGTQ 110
           |  | : |  |||:|:||| : | :| ||||| |||| || :||::|:||: | ||: |
Qy      13 FSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 72

Db     111 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGFYWPEMLKCDKFP-EGDVC 166
            ||||||||||||     | ||: ||  |:| | |||  ||| ||:|| ||:|| : |:|
Qy      73 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 132

Db     167 IAMTPPNPTEASKPQGTTVCPPCDNELKSEA-IIEHLCASEFALRMKIKEVKKENGDKKI 225
           |::::::   ::  ::  || :| |    :  |:| || ::|||::|:||:   | | ||
Qy     133 IPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKI 192

Db     226 V-PKKKKPL-KLGPIKKKDLKKLVLYLKNGADCPCHQLDNLSHHFLIMGRKVKSQYLLTA 283
           :   | | : ||  :  :|||| || ||:: :| | :::::  :|:||:|  :: ::|:
Qy     193 ILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 252

Db     284 IHKWDKKNKEFKNFMKKMKNHEC 306
           : :|:|  :||| : : ::: :|
Qy     253 VKRWQKGQREFKRISRSIRKLQC 275


RESULT    6
ID   O00546      PRELIMINARY;      PRT;   313 AA.
AC   O00546;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED FRIZZLED-RELATED PROTEIN.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   FINCH P.W., HE X., KELLEY M.J., UREN A., SCHAUDIES R.P., POPESCU N.C.,
RA   RUDIKOFF S., AARONSON S.A., VARMUS H.E., RUBIN J.S.;
RL   SUBMITTED (JUN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   ZHOU Z., WANG J.;
RL   SUBMITTED (MAR-1998) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF001900; G2213819; -.
DR   EMBL; AF056087; G3033551; -.
SQ   SEQUENCE   313 AA;  35314 MW;  A3CD04FF CRC32;

  Query Match           43.9%;  Score 911;  DB 4;  Length 313;
  Best Local Similarity 42.6%;  Pred. No. 9.94e-193;
  Matches  112;  Conservative   67;  Mismatches  77;  Indels   7;  Gaps   5;

Db      50 FYTKPPQCVDIPADLRLCHNVGYKKMVLPNLLEHETMAEVKQQASSWVPLLNKNCHAGTQ 109
           |  | : |  |||:|:||| : | :| ||||| |||| || :||::|:||: | ||: |
Qy      13 FSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 72

Db     110 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGFYWPEMLKCDKFP-EGDVC 165
            ||||||||||||     | ||: ||  |:| | |||  ||| ||:|| ||:|| : |:|
Qy      73 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 132

Db     166 IAMTPPNATEASKPQGTTVCPPCDNELKSEA-IIEHLCASEFALRMKIKEVKKENGDKKI 224
           |::::::   ::  ::  || :| |    :  |:| || ::|||::|:||:   | | ||
Qy     133 IPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKI 192

Db     225 V-PKKKKPL-KLGPIKKKDLKKLVLYLKNGADCPCHQLDNLSHHFLIMGRKVKSQYLLTA 282
           :   | | : ||  :  :|||| || ||:: :| | :::::  :|:||:|  :: ::|:
Qy     193 ILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 252

Db     283 IHKWDKKNKEFKNFMKKMKNHEC 305
           : :|:|  :||| : : ::: :|
Qy     253 VKRWQKGQREFKRISRSIRKLQC 275


RESULT    7
ID   O19116      PRELIMINARY;      PRT;   308 AA.
AC   O19116;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   FRZA PRECURSOR.
OS   BOS TAURUS (BOVINE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RA   DUPLAA C., D'AMORE P.A.;
RL   SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U85945; G2337937; -.
KW   SIGNAL.
FT   SIGNAL        5      22       POTENTIAL.
SQ   SEQUENCE   308 AA;  34763 MW;  9B74A92F CRC32;

  Query Match           43.5%;  Score 902;  DB 6;  Length 308;
  Best Local Similarity 42.2%;  Pred. No. 1.91e-190;
  Matches  111;  Conservative   66;  Mismatches  79;  Indels   7;  Gaps   5;

Db      45 FYTKPPQCVDIPADLRLCHNVGYKRMVLPNLLEHETMAEVKQQASSWVPLLNKNCHIGTQ 104
           |  | : |  |||:|:||| : |  | |||| |||| || :||::|:||: | ||  |
Qy      13 FSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 72

Db     105 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGFYWPEMLKCDKFP-EGDVC 160
            ||||||||||||     | ||: ||  |:| | |||  ||| ||:|| ||:|| : |:|
Qy      73 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 132

Db     161 IAMTPPNATEASKPQGTTVCPPCDNELKSEA-IIEHLCASEFALRMKIKEVKKENGDKKI 219
           |::::::   ::  ::  || :| |    :  |:| || ::|||::|:||:   | | ||
Qy     133 IPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKI 192

Db     220 V-PKKKKPL-KLGPIKKKELKKLVLYLKNGADCPCHQLDNLSHHFLIMGRKVKSQYLLTA 277
           :   | | : ||  :  ::||| || ||:: :| | :::::  :|:||:|  :: ::|:
Qy     193 ILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 252

Db     278 IHKWDKKNKEFKTFMKKMKNHEC 300
           : :|:|  :||| : : ::: :|
Qy     253 VKRWQKGQREFKRISRSIRKLQC 275


RESULT    8
ID   O14780      PRELIMINARY;      PRT;   317 AA.
AC   O14780;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED APOPTOSIS RELATED PROTEIN 3.
GN   SARP3.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PANCREAS;
RA   MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M.,
RA   FITZPATRIC P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF017988; G2415419; -.
SQ   SEQUENCE   317 AA;  35577 MW;  2CD69373 CRC32;

  Query Match           43.5%;  Score 901;  DB 4;  Length 317;
  Best Local Similarity 45.2%;  Pred. No. 3.42e-190;
```

```
  Matches   114;  Conservative   58;  Mismatches  74;  Indels   6;  Gaps   4;

Db      46 SYSKPPQCLDIPADLPLCHTVGYKRMRLPNLLEHESLAEVKQQASSWLPLLAKRCHSDTQ 105
           || | : |  |||:| ||| : |  ||||||| ||:: || :||::|:||: |:||:||
Qy      14 SY-KRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 72

Db     106 VFLCSLFAPVCLD---RPIYPCRSLCEAVRAGCAPLMEAYGFPWPEMLHCHKFPLDNDLC 162
            ||||||||||||     | ||:|||  |:  |||:| |:||||||:|| | :|| ||||
Qy      73 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 132

Db     163 IAVQFG-HLPATAPPVTKICAQC-EMEHSADGLMEQMCSSDFVVKMRIKEIKIENGDRKL 220
           |::  : || :::    |:|  | :::   : :|| :| :|| :|:::|||   | | |:
Qy     133 IPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKI 192

Db     221 IGAQKKKKLLKPGPLKRKDTKRLVLHMKNGAGCPCPQLDSLAGSFLVMGRKVDGQLLLMA 280
           |   | | : |   :  :| |: || :|::  | | ::: : :::|||||:|  |:|:: :
Qy     193 ILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 252

Db     281 VYRWDKKNKEMK 292
           | ||:|  :| |
Qy     253 VKRWQKGQREFK 264


RESULT    9
ID   O08861      PRELIMINARY;      PRT;   314 AA.
AC   O08861;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   SECRETED FRIZZLED RELATED PROTEIN SFRP-1.
GN   SFRP1.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97250455.
RA   RATTNER A., HSIEH J.C., SMALLWOOD P.M., GILBERT D.J., COPELAND N.G.,
RA   JENKINS N.A., NATHANS J.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 94:2859-2863(1997).
DR   EMBL; U88566; G1946341; -.
SQ   SEQUENCE   314 AA;  35327 MW;  5AA5CA27 CRC32;

  Query Match           42.9%;  Score 890;  DB 11;  Length 314;
  Best Local Similarity  41.4%;  Pred. No. 2.10e-187;
  Matches   109;  Conservative   67;  Mismatches  80;  Indels   7;  Gaps   5;

Db      51 FYTKPPQCVDIPVDLRLCHNVGYKKMVLPNLLEHETMAEVKQQASSWVPLLNKNCHMGTQ 110
           |  | : |  || :|:||| : | :| ||||| |||| || :||::|:||: | ||  |
Qy      13 FSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 72

Db     111 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGFYWPEMLKCDKFP-EGDVC 166
            ||||||||||||     | ||: ||  |:| | |||  ||| ||:|| ||:|| : |:|
Qy      73 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 132

Db     167 IAMTPPNTTEASKPQGTTVCPPCDNELKSEA-IIEHLCASEFALRMKIKEVKKENGDKKI 225
           |::::::   ::  ::  || :| |    :  |:| || ::|||::|:|||:   | | ||
Qy     133 IPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKI 192

Db     226 V-PKKKKPL-KLGPIKKKELKALVLFLKNGADCPCHQLDNLSHNFLIMGRKVKSQYLLTA 283
           :   | | : ||  :  ::||  || ||:: :| | ::::::  :|:||:|  :: ::|:
Qy     193 ILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 252

Db     284 IHKWDKKNKEFKNFMKRMKNHEC 306
           : :|:|  :||| : : ::: :|
Qy     253 VKRWQKGQREFKRISRSIRKLQC 275


RESULT   10
ID   O42397      PRELIMINARY;      PRT;   307 AA.
AC   O42397;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   CRESCENT.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   PFEFFER P.L., IZPISUA-BELMONTE J.C., DE ROBERTIS E.M.;
RL   SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF006508; G2226372; -.
SQ   SEQUENCE   307 AA;  34683 MW;  320A2243 CRC32;

  Query Match           35.2%;  Score 729;  DB 13;  Length 307;
  Best Local Similarity  35.8%;  Pred. No. 8.17e-147;
  Matches    96;  Conservative   68;  Mismatches  94;  Indels  10;  Gaps   8;

Db      36 SYLRRSSSCTAIPRSMALCYDIGYSEMRIPNLLEHETMPEVIQQSSSWLPLLARECHPDA 95
           || :||: | :|| :: ||  | | :||:|||| |||| ||::|:::|:||: ::||||:
Qy      14 SY-KRSN-CKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDT 71

Db      96 RIFLCSLFAPICLDRL---IYPCRSLCEAVKRSCAPVMACYGYPWPEILNCNKFPADHEL 152
           : |||||||:||| |   | ||:|||  ||  |||||: :|:||||::|:|::||  |::|
Qy      72 KKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDL 131

Db     153 CIAAVSTDES-SSSRRMPRASCKDCELEEASTAREILDNLCANDFTVKIRILRKNTTTTV 211
           ||:  |:|   :::   |:  |  |  :      :|:: || |||::||::
Qy     132 CIPLASSDHLLPATEEAPKV-CEACK-NKNDDDNDIMETLCKNDFALKIKVKEITYINRD 189

Db     212 SDFDLDPSRVEVLKPGPLLRTEIPGRLQQWLDIDATCAHNIMRGTHAGVFVISGEVRSDK 271
           : : |:     : |   :   ::   :  ||     |: : |   :|  ::: |: ::
Qy     190 TKIILETKSKTIYKLNGVSERDLKKSVL-WLKDSLQCTCEEMNDINAP-YLVMGQKQGGE 247

Db     272 VVVNKAYAWQKKNRNLHQAVRRWKHHRC 299
           :|:     |||  |:: :  |  :  :|
Qy     248 LVITSVKRWQKGQREFKRISRSIRKLQC 275


RESULT   11
ID   O73821      PRELIMINARY;      PRT;   281 AA.
AC   O73821;
DT   01-AUG-1998 (TREMBLREL. 07, CREATED)
DT   01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED XWNT8 INHIBITOR SIZZLED.
GN   SZL.
OS   XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 98088686.
RA   SALIC A.N., KROLL K.L., EVANS L.M., KIRSCHNER M.W.;
RL   DEVELOPMENT 124:4739-4748(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   SALIC A.N., KROLL K.L., EVANS L.M., KIRSCHNER M.W.;
RL   SUBMITTED (APR-1998) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF059570; G3089551; -.
SQ   SEQUENCE   281 AA;  31834 MW;  8478D86B CRC32;

  Query Match           28.1%;  Score 582;  DB 13;  Length 281;
  Best Local Similarity  32.7%;  Pred. No. 3.03e-110;
  Matches    86;  Conservative   65;  Mismatches 101;  Indels  11;  Gaps   8;

Db      25 TKCVTIPTEMAMCNDVGYSEMRLPNLMGHTNMAEVVPKSAEWQNLLQTGCHPYARTFLCS 84
           ::|  ||::: :|: : | :|||||:||  | ||:  :: |  |:   ||| :: ||||
Qy      18 SNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCS 77

Db      85 LFAPVCLDTF---IQPCRSMCVAVRDSCAPVLACHGHSWPESLDCDRFPAGEDMCLDTLS 141
           |||||||| :   ||||:|:|| |:| ||||::  | :||: |:|||||  :|:|:   |
Qy      78 LFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLAS 137

Db     142 KEYQY-SYKELPKPSCQGCPLIEEFFSHKTVLEAFCDNNFAVKVKLAKKKSAS-GL-YEY 198
            :    : | ||  |::|   ::    : ::|::| |:||:|:|:      :
```

```
Qy    138 SDHLLPATEEAPK-VCEACKNKND--DDNDIMETLCKNDFALKIKVKEITYINRDTKIIL 194

Db    199 ETEGPVEFIKQGLLLPYDTRTMIEQWLLINENCAQKLIRTRPTVYVIAGEIHHGKVKVNR 258
Qy    195 ETKSKTIYKLNGVSER-DLKKSV-LWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 252

Db    259 IFHWQKKDSQLTLATRRWRHHKC 281
Qy    253 VKRWQKGQREFKRISRSIRKLQC 275


RESULT   12
ID   O16147      PRELIMINARY;      PRT;   568 AA.
AC   O16147;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   MOM-5.
GN   MOM-5.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL;
RX   MEDLINE; 97433081.
RA   ROCHELEAU C.E., DOWNS W.D., LIN R., WITTMANN C., BEI Y., CHA Y.-H.,
RA   ALI M., PRIESS J.R., MELLO C.C.;
RL   CELL 90:707-716(1997).
DR   EMBL; AF013953; G2463674; -.
SQ   SEQUENCE   568 AA;  62892 MW;  8578BB4F CRC32;

  Query Match           19.2%;  Score 398;  DB 5;  Length 568;
  Best Local Similarity  40.2%;  Pred. No. 1.38e-65;
  Matches    51;  Conservative   25;  Mismatches  46;  Indels   5;  Gaps   5;

Db     31 FSTTR-KCEHI-T-IPMCKNLDYNQTVFPNLLGHTTQSEAGPAIAQFNPLIKVKCSEDIR 87
Qy     13 FSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 72

Db     88 LFLCTVYAPVC-TVLEKPIQPCRELCLSAKNGCESLMKKFGFQWPDQLDCNKFPV-TDLC 145
Qy     73 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 132

Db    146 VGKNSSE 152
Qy    133 IPLASSD 139


RESULT   13
ID   O57328      PRELIMINARY;      PRT;   592 AA.
AC   O57328;
DT   01-JUN-1998 (TREMBLREL. 06, CREATED)
DT   01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   7-TRANSMEMBRANE PROTEIN FRIZZLED-1.
GN   FZ-1.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIMB BUD;
RA   KENGAKU M., TWOMBLY V., TABIN C.;
RL   COLD SPRING HARB. SYMP. QUANT. BIOL. 0:0-0(1997).
DR   EMBL; AF031830; G2655274; -.
KW   TRANSMEMBRANE.
SQ   SEQUENCE   592 AA;  65490 MW;  CD66EC83 CRC32;

  Query Match           19.2%;  Score 397;  DB 13;  Length 592;
  Best Local Similarity  37.7%;  Pred. No. 2.38e-65;
  Matches    49;  Conservative   32;  Mismatches  44;  Indels   5;  Gaps   5;

Db     67 HGYCQPI-S-IPLCTDIAYNQTIMPNLLGHTNQEDAGLEVHQFYPLVKVQCSAELKFFLC 124
Qy     17 RSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 76

Db    125 SMYAPVC-TVLEQALPPCRSLCERARQGCEALMNKFGFQWPDTLRCEKFPVHGAGELCVG 183
Qy     77 SLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QDN-DLCIP 134

Db    184 QNASERGTPT 193
Qy    135 LASSDHLLPA 144


RESULT   14
ID   Q61090      PRELIMINARY;      PRT;   572 AA.
AC   Q61090;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   FRIZZLED HOMOLOG 7 (TRANSMEMBRANE RECEPTOR).
GN   FZD7 OR FRIZZLED 7.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96224032.
RA   WANG Y., MACKE J.P., ABELLA B.S., ANDREASSON K., WORLEY P.,
RA   GILBERT D.J., COPELAND N.G., JENKINS N.A., NATHANS J.;
RL   J. BIOL. CHEM. 271:4468-4476(1996).
DR   EMBL; U43320; G1151258; -.
DR   MGD; MGI:108570; FZD7.
KW   TRANSMEMBRANE.
SQ   SEQUENCE   572 AA;  63816 MW;  FA1D78D2 CRC32;

  Query Match           19.1%;  Score 396;  DB 11;  Length 572;
  Best Local Similarity  40.7%;  Pred. No. 4.11e-65;
  Matches    48;  Conservative   25;  Mismatches  42;  Indels   3;  Gaps   3;

Db     51 PISIPLCTDIAYNQTILPNLLGHTNQEDAGLEVHQFYPLVKVQCSPELRFFLCSMYAPVC 110
Qy     24 PANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVC 83

Db    111 -TVLDQAIPPCRSLCERARQGCEALMNKFGFQWPERLRCENFPVHGAGEICVGQNTSD 167
Qy     84 LDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QDN-DLCIPLASSD 139


RESULT   15
ID   O00144      PRELIMINARY;      PRT;   591 AA.
AC   O00144;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   FRIZZLED HOMOLOG.
GN   FZD3.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97227293.
RA   WANG Y.K., HARRYMAN SAMOS C., PEOPLES R., PEREZ-JURADO L.A., NUSSE R.,
RA   FRANCKE U.;
RL   HUM. MOL. GENET. 6:465-472(1997).
DR   EMBL; U82169; G1906598; -.
SQ   SEQUENCE   591 AA;  64466 MW;  88C5EC1F CRC32;

  Query Match           19.0%;  Score 394;  DB 4;  Length 591;
  Best Local Similarity  37.5%;  Pred. No. 1.23e-64;
  Matches    48;  Conservative   32;  Mismatches  47;  Indels   1;  Gaps   1;

Db     41 AVEIPMCRGIGYNLTRMPNLLGHTSQGEAAAELAEFAPLVQYGCHSHLRFFLCSLYAPMC 100
```

```
Qy      24 PANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVC 83

Db     101 TDQVSTPIPACRPMCEQARLRCAPIMEQFNFGWPDSLDCARLPTRNDPHALCMEAPENAT 160
           |::   | :|:::| | : ||||:|  | | ||| |:| |:|  ||   : : ::::
Qy      84 LDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDL-CIPLASSDHLL 142

Db     161 AGPAEPHK 168
           ::  |: |
Qy     143 PATEEAPK 150


Search completed: Thu Oct 22 15:58:52 1998
Job time : 85 secs.
```

```
**************************************************************************
MPSRCH (TM)
**************************************************************************
```



```
MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Thu Oct 22 15:59:10 1998;  MasPar time 12.71 Seconds
                                            790.336 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-848-439-3
Description:     (1-275) from US08848439.pep
Perfect Score:   2073
Sequence:        1 SARGLFLFGQPDFSYKRSNC..........WQKGQREFKRISRSIRKLQC 275

Scoring table:   PAM 150
                 Gap 11

Searched:        120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        pir56
                   1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d

Statistics:      Mean 44.556;  Variance 87.725;  scale 0.508

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	386	18.6	641	2	A45054	probable intercellula	1.05e-51
2	355	17.1	694	2	S78444	dfz2 protein - fruit	1.19e-45
3	355	17.1	694	2	S71786	dfz2 protein - fruit	1.19e-45
4	332	16.0	415	2	S15709	hypothetical protein	3.34e-41
5	332	16.0	581	2	S03540	gene frizzled protein	3.34e-41
6	178	8.6	562	2	B56101	collagen alpha 1(XVII	4.96e-13
7	148	7.1	581	2	B54665	netrin-2 precursor -	4.04e-08
8	124	6.0	606	2	A54665	netrin-1 precursor -	1.87e-04
9	122	5.9	793	2	JC5539	Smoothened protein -	3.66e-04
10	101	4.9	1735	2	S22812	DNA-directed RNA poly	2.94e-01
11	99	4.8	322	2	H64825	hypothetical protein	5.34e-01
12	100	4.8	500	2	D31579	cytochrome P450 2D4 -	3.97e-01
13	100	4.8	500	2	I52313	cytochrome P450 2D4 h	3.97e-01
14	99	4.8	724	2	B32571	ribosomal protein S6	5.34e-01
15	98	4.7	391	2	E64474	hypothetical protein	7.17e-01
16	98	4.7	1032	1	G7BPT4	baseplate protein gp7	7.17e-01
17	97	4.7	15281	2	S41309	cyclosporin synthetas	9.62e-01
18	96	4.6	520	2	JS0291	intermediate filament	1.29e+00
19	96	4.6	735	2	I51901	ribosomal protein S6	1.29e+00
20	94	4.5	271	2	E64694	hypothetical protein	2.29e+00
21	93	4.5	735	2	A53300	ribosomal protein S6	3.04e+00
22	92	4.4	203	2	B70254	conserved hypothetica	4.03e+00
23	92	4.4	608	2	B70188	hypothetical protein	4.03e+00
24	92	4.4	801	2	B64679	paralysed flagella pr	4.03e+00
25	92	4.4	1562	2	S53069	probable membrane pro	4.03e+00
26	90	4.3	176	2	S70008	troponin I - Atlantic	7.04e+00
27	89	4.3	262	2	A60105	virF protein - Shigel	9.27e+00
28	89	4.3	262	2	A47605	virF virulence protei	9.27e+00
29	89	4.3	262	2	S14646	virF protein - Shigel	9.27e+00
30	90	4.3	263	2	S73489	S-adenosylmethionine-	7.04e+00
31	89	4.3	508	2	A64571	GMP synthase - Helico	9.27e+00
32	90	4.3	521	2	A53153	glucose transport pro	7.04e+00
33	89	4.3	752	2	A32571	ribosomal protein S6	9.27e+00
34	88	4.2	127	2	S45967	probable membrane pro	1.22e+01
35	88	4.2	182	2	S28310	hypothetical protein	1.22e+01
36	88	4.2	183	2	D69108	phycocyanin alpha phy	1.22e+01
37	88	4.2	370	2	A69430	conserved hypothetica	1.22e+01
38	88	4.2	443	2	JQ1527	alpha-amylase (EC 3.2	1.22e+01
39	88	4.2	445	2	S19990	alpha-amylase (EC 3.2	1.22e+01
40	88	4.2	633	2	C32571	ribosomal protein S6	1.22e+01
41	88	4.2	740	2	I38556	insulin-stimulated pr	1.22e+01
42	88	4.2	849	2	S00030	neurofilament triplet	1.22e+01
43	88	4.2	871	2	I48696	gene Nsk2 protein - m	1.22e+01
44	88	4.2	881	2	I48697	gene Nsk2 protein - m	1.22e+01
45	88	4.2	1102	2	S44772	C29E4.4 protein - Cae	1.22e+01

ALIGNMENTS

```
RESULT      1
ENTRY             A45054     #type complete
TITLE             probable intercellular signal transducer or transmitter Fz-1
                    - rat
ORGANISM          #formal_name Rattus norvegicus #common_name Norway rat
DATE              30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change
                    28-Apr-1995
ACCESSIONS        A45054
REFERENCE         A45054
   #authors       Chan, S.D.; Karpf, D.B.; Fowlkes, M.E.; Hooks, M.; Bradley,
                    M.S.; Vuong, V.; Bambino, T.; Liu, M.Y.; Arnaud, C.D.;
                    Strewler, G.J.; Nissenson, R.A.
   #journal       J. Biol. Chem. (1992) 267:25202-25207
   #title         Two homologs of the Drosophila polarity gene frizzled (fz)
                    are widely expressed in mammalian tissues.
   #cross-references MUID:93094228
   #accession     A45054
      ##status       preliminary; not compared with conceptual translation
      ##molecule_type nucleic acid
      ##residues        1-641 ##label CHA
      ##experimental_source UMR 106 osteosarcoma cell line
      ##note            sequence extracted from NCBI backbone (NCBIP:120154)
SUMMARY           #length 641  #molecular-weight 71054  #checksum 8376

  Query Match             18.6%;  Score 386;  DB 2;  Length 641;
  Best Local Similarity   39.2%;  Pred. No. 1.05e-51;
  Matches    49;  Conservative   27;  Mismatches   43;  Indels    6;  Gaps    6;

Db     108 HGYCQPI-S-IPLCTDIAYNQTIMPNLLGHTNQEDAGLEVHQFYPLVKVQCSAELKFFLC 165
           :: | || : : || | |    :||||||    :    :      |||: || :: | |||
Qy      17 RSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 76

Db     166 SMYAPVC-TVLEQALPPCRSLCERAQG-CEALMNKFGFQWPDTLKCEKFPVHGRGELCVG 223
           |::||||   |:::: ||:|||  :     | ::|: ||| ||| | |::|| :    :||:
Qy      77 SLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QD-NDLCIP 134

Db     224 QNTSD 228
            :||
Qy     135 LASSD 139


RESULT      2
ENTRY             S78444     #type complete
TITLE             dfz2 protein - fruit fly (Drosophila melanogaster)
ORGANISM          #formal_name Drosophila melanogaster
DATE              12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change
                    13-Mar-1998
```

```
ACCESSIONS       S78444
REFERENCE        S78444
   #authors      Bhanot, P.; Wang, Y.; Nathans, J.
   #submission   submitted to the EMBL Data Library, July 1996
   #accession    S78444
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-694 ##label BHA
      ##cross-references EMBL:U65589
SUMMARY          #length 694  #molecular-weight 75437  #checksum 3113

  Query Match            17.1%;  Score 355;  DB 2;  Length 694;
  Best Local Similarity  36.2%;  Pred. No. 1.19e-45;
  Matches    47;  Conservative   32;  Mismatches  45;  Indels   6;  Gaps   6;

Db      51 YGVPAIP-KDPNLRCEEITIPMCRGIGYNMTSFPNEMNHETQDEAGLEVHQFW-PLVEIK 108
           :| | :: | :| :     : :|:|| |    :|| : |||  |  ||    | |||
Qy       8 FGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQ 66

Db     109 CSPDLKFFLCSMYTPICLEDYHKPLPVCRSVCERARSGCAPIMQQYSFEWPERMACEHLP 168
           | || | ||||:::|:||:|    :  |:|:| : :  |||:|  ::| ||: : |:::|
Qy      67 CHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP 126

Db     169 LHGDPDNLCM 178
            : | ::||:
Qy     127 -Q-D-NDLCI 133


RESULT    3
ENTRY            S71786     #type complete
TITLE            dfz2 protein - fruit fly (Drosophila melanogaster)
ORGANISM         #formal_name Drosophila melanogaster
DATE             12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change
                   13-Mar-1998
ACCESSIONS       S71786
REFERENCE        S71786
   #authors      Bhanot, P.; Brink, M.; Harryman Samos, C.; Hsieh, J.C.; Wang,
                   Y.; Macke, J.P.; Andrew, D.; Nathans, J.; Nusse, R.
   #journal      Nature (1996) 382:225-230
   #title        A new member of the frizzled family from Drosophila functions
                   as a wingless receptor.
   #accession    S71786
      ##status        preliminary; nucleic acid sequence not shown
      ##molecule_type DNA
      ##residues      1-694 ##label BHA
      ##cross-references EMBL:U65589
SUMMARY          #length 694  #molecular-weight 75423  #checksum 3095

  Query Match            17.1%;  Score 355;  DB 2;  Length 694;
  Best Local Similarity  36.2%;  Pred. No. 1.19e-45;
  Matches    47;  Conservative   32;  Mismatches  45;  Indels   6;  Gaps   6;

Db      51 YGVPAIP-KDPNLRCEEITIPMCRGIGYNMTSFPNEMNHETQDEAGLEVHQFW-PLVEIK 108
           :| | :: | :| :     : :|:|| |    :|| : |||  |  ||    | |||
Qy       8 FGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQ 66

Db     109 CSPDLKFFLCSMYTPICLEDYHKPLPVCRSVCERARSGCAPIMQQYSFEWPERMACEHLP 168
           | || | ||||:::|:||:|    :  |:|:| : :  |||:|  ::| ||: : |:::|
Qy      67 CHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP 126

Db     169 LHGDPDNLCM 178
            : | ::||:
Qy     127 -Q-D-NDLCI 133


RESULT    4
ENTRY            S15709     #type complete
TITLE            hypothetical protein - fruit fly (Drosophila melanogaster)
ORGANISM         #formal_name Drosophila melanogaster
DATE             21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
                   16-Feb-1997
ACCESSIONS       S15709
REFERENCE        S15708
   #authors      Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
   #journal      Genetics (1990) 126:401-416
   #title        Molecular structure of frizzled, a Drosophila tissue polarity
                   gene.
   #cross-references MUID:91060073
   #accession    S15709
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-415 ##label ADL
      ##cross-references EMBL:X54648
GENETICS
   #gene         FlyBase:fz
      ##cross-references FlyBase:FBgn0001085
   #introns      224/3; 264/3; 329/3; 405/1
SUMMARY          #length 415  #molecular-weight 46117  #checksum 4401

  Query Match            16.0%;  Score 332;  DB 2;  Length 415;
  Best Local Similarity  37.5%;  Pred. No. 3.34e-41;
  Matches    45;  Conservative   22;  Mismatches  47;  Indels   6;  Gaps   6;

Db      50 HNRCEPI-T-ISICKNIPYNMTIMPNLIGHTKQEEAGLEVHQFAPLVKIGCSDDLQLFLC 107
           :: | || : : :|  | |    :|||:||    |   :     |||:  |  |   |||
Qy      17 RSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 76

Db     108 SLYVPVC-TILERPIPPCRSLCESARV-CEKLMKTYNFNWPENLECSKFPVHGGEDLCVA 165
           ||: |||   |:  | ||:|||   :  |  :| :: | ||: ||| :|| :  :|||::
Qy      77 SLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QD-NDLCIP 134


RESULT    5
ENTRY            S03540     #type complete
TITLE            gene frizzled protein precursor - fruit fly (Drosophila
                   melanogaster)
ORGANISM         #formal_name Drosophila melanogaster
DATE             31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
                   23-Aug-1997
ACCESSIONS       S03540; S15708
REFERENCE        S03540
   #authors      Vinson, C.R.; Conover, S.; Adler, P.N.
   #journal      Nature (1989) 338:263-264
   #title        A Drosophila tissue polarity locus encodes a protein
                   containing seven potential transmembrane domains.
   #cross-references MUID:89159415
   #accession    S03540
      ##status        not compared with conceptual translation
      ##molecule_type mRNA
      ##residues      1-581 ##label VIN
REFERENCE        S15708
   #authors      Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
   #journal      Genetics (1990) 126:401-416
   #title        Molecular structure of frizzled, a Drosophila tissue polarity
                   gene.
   #cross-references MUID:91060073
   #accession    S15708
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-581 ##label ADL
      ##cross-references EMBL:X54648
GENETICS
   #gene         FlyBase:fz
      ##cross-references FlyBase:FBgn0001085
   #introns      224/3; 264/3; 329/3; 405/1
KEYWORDS         alternative splicing; transmembrane protein
FEATURE
   1-26               #domain signal sequence #status predicted #label SIG\
   27-581             #product gene frizzled protein #status predicted #label
                        MAT
SUMMARY          #length 581  #molecular-weight 64846  #checksum 358

  Query Match            16.0%;  Score 332;  DB 2;  Length 581;
  Best Local Similarity  37.5%;  Pred. No. 3.34e-41;
  Matches    45;  Conservative   22;  Mismatches  47;  Indels   6;  Gaps   6;
```

```
Db      50 HNRCEPI-T-ISICKNIPYNMTIMPNLIGHTKQEEAGLEVHQFAPLVKIGCSDDLQLFLC 107
           :: | || : : :|  | |    :|||:||    |   :     |||:  |  |   |||
Qy      17 RSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 76

Db     108 SLYVPVC-TILERPIPPCRSLCESARV-CEKLMKTYNFNWPENLECSKFPVHGGEDLCVA 165
           ||: |||   |:  | ||:|||   :  |  :| :: | ||: ||| :|| :  :|||::
Qy      77 SLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QD-NDLCIP 134


RESULT    6
ENTRY            B56101     #type fragment
TITLE            collagen alpha 1(XVIII) chain precursor long form - mouse
                   (fragment)
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change
                   10-Sep-1997
ACCESSIONS       B56101
REFERENCE        A56101
   #authors      Rehn, M.; Pihlajaniemi, T.
   #journal      J. Biol. Chem. (1995) 270:4705-4711
   #title        Identification of three N-terminal ends of type XVIII
                   collagen chains and tissue-specific differences in the
                   expression of the corresponding transcripts. The longest
                   form contains a novel motif homologous to rat and
                   Drosophila frizzled proteins.
   #accession    B56101
      ##status       preliminary
      ##molecule_type mRNA
      ##residues      1-562 ##label REH
      ##cross-references GB:U11637; NID:g618429; PID:g618430
GENETICS
   #gene         COL18A1
SUMMARY          #length 562  #checksum 115

  Query Match             8.6%;  Score 178;  DB 2;  Length 562;
  Best Local Similarity  34.1%;  Pred. No. 4.96e-13;
  Matches    30;  Conservative   13;  Mismatches  45;  Indels   0;  Gaps   0;

Db     368 SRCLPLPPTLTLCSRLGIGHFWLPNHLHHTDSVEVEATVQAWGRFLHTNCHPFLAWFFCL 427
           | | |:|: | ||  :   : :||| | |    ||     ||  ::   |||    |:|
Qy      18 SNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCS 77

Db     428 LLAPSCGPGPPPPLPPCRQFCEALEDEC 455
           |:|| |       : ||: :|  : | |
Qy      78 LFAPVCLDDLDETIQPCHSLCVQVKDRC 105


RESULT    7
ENTRY            B54665     #type fragment
TITLE            netrin-2 precursor - chicken (fragment)
ORGANISM         #formal_name Gallus gallus #common_name chicken
DATE             06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
                   10-Sep-1997
ACCESSIONS       B54665
REFERENCE        A54665
   #authors      Serafini, T.; Kennedy, T.E.; Galko, M.J.; Mirzayan, C.;
                   Jessell, T.M.; Tessier-Lavigne, M.
   #journal      Cell (1994) 78:409-424
   #title        The netrins define a family of axon outgrowth-promoting
                   proteins homologous to Caenorhabditis elegans UNC-6.
   #accession    B54665
      ##status       preliminary
      ##molecule_type mRNA
      ##residues      1-581 ##label SER
      ##cross-references GB:L34550; NID:g529420; PID:g529421
SUMMARY          #length 581  #checksum 7731

  Query Match             7.1%;  Score 148;  DB 2;  Length 581;
  Best Local Similarity  27.0%;  Pred. No. 4.04e-08;
  Matches    30;  Conservative   28;  Mismatches  46;  Indels   7;  Gaps   6;

Db     431 IPAINPTSLVTSTE-APADCDSYCKPAKGNYKINMKKYCKKDYVVQVNILEMETVANWAK 489
           ||  ::  |: :|| ||  |:: ||  : : :  |   ||:|: : ::: |:  :   :|
Qy     133 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTK 191

Db     490 FTINILS--VYKCRDERVKR-GDNFLWIHLKDLSCKCPKI-QISKKYLVMG 536
           : ::  |  :||      :    : ||:    | | |  : :|:  |||||
Qy     192 IILETKSKTIYKLNGVSERDLKKSVLWLK-DSLQCTCEEMNDINAPYLVMG 241


RESULT    8
ENTRY            A54665     #type complete
TITLE            netrin-1 precursor - chicken
ORGANISM         #formal_name Gallus gallus #common_name chicken
DATE             06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
                   10-Sep-1997
ACCESSIONS       A54665
REFERENCE        A54665
   #authors      Serafini, T.; Kennedy, T.E.; Galko, M.J.; Mirzayan, C.;
                   Jessell, T.M.; Tessier-Lavigne, M.
   #journal      Cell (1994) 78:409-424
   #title        The netrins define a family of axon outgrowth-promoting
                   proteins homologous to Caenorhabditis elegans UNC-6.
   #accession    A54665
      ##status       preliminary
      ##molecule_type mRNA
      ##residues      1-606 ##label SER
      ##cross-references GB:L34549; NID:g529418; PID:g529419
SUMMARY          #length 606  #molecular-weight 68126  #checksum 5400

  Query Match             6.0%;  Score 124;  DB 2;  Length 606;
  Best Local Similarity  26.1%;  Pred. No. 1.87e-04;
  Matches    31;  Conservative   34;  Mismatches  41;  Indels  13;  Gaps  11;

Db     456 IPAAPPPTAASSTEEPAD-CDSYCKASKGKLKINMKKYCKKDYAVQIHILKAEKNADW-W 513
           || |::    ::|||::  |:: ||  :   :  |   ||:|:|: | : |     ::
Qy     133 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDT 190

Db     514 KFTVNIIS--VYK-QG-SNR-LRRGDQTLWVHAKDIACKCPKVKPMKK-YLLLGSTEDS 566
           |: ::  |  :||  | |:| |:::   ||:    : | |  :: ::  ||::|  : :
Qy     191 KIILETKSKTIYKLNGVSERDLKKS--VLWLKD-SLQCTCEEMNDINAPYLVMGQKQGG 246


RESULT    9
ENTRY            JC5539     #type complete
TITLE            Smoothened protein - mouse
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change
                   05-Sep-1997
ACCESSIONS       JC5539; PC4476
REFERENCE        JC5539
   #authors      Akiyama, H.; Shigeno, C.; Hiraki, Y.; Shukunami, C.; Kohno,
                   H.; Akagi, M.; Konishi, J.; Nakamura, T.
   #journal      Biochem. Biophys. Res. Commun. (1997) 235:142-147
   #title        Cloning of a mouse Smoothened cDNA and expression patterns of
                   hedgehog signalling molecules during chondrogenesis and
                   cartilage differentiation in clonal mouse EC cells, ATDC5.
   #accession    JC5539
      ##molecule_type mRNA
      ##residues      1-793 ##label AKI
   #accession    PC4476
      ##molecule_type protein
      ##residues      528-533;539-545;600-605 ##label AK2
      ##experimental_source ADTC5 cell
COMMENT     This protein is used in the conserved targets in Hedgehog
              signalling pathway, together with Patched and Gli. These protein
              are responsible for the skeletal abnormalities in Gorlin and
              Greig syndromes.
FEATURE
   1-32              #domain signal sequence #status predicted #label SIG
SUMMARY          #length 793  #molecular-weight 87299  #checksum 8609

  Query Match             5.9%;  Score 122;  DB 2;  Length 793;
  Best Local Similarity  26.7%;  Pred. No. 3.66e-04;
  Matches    31;  Conservative   22;  Mismatches  54;  Indels   9;  Gaps   7;
```

```
Db      70 RAAHCEPLRYNVCLGSALPYGATT-TLLAGDSDSQEEAHGKLVLWSGLRNAPRCWAVIQP 128
           : ::| |:  |: |  :: |       | |  ::  |       |  |    :| :
Qy      16 KRSNCKPIPANLQLCHGIEYQNMRLPNLLGH-ETMKEVLEQAGAWIPL-VMKQCHPDTKK 73

Db     129 LLCAVYMPKC-EN-DRVELPSRTLCQATRGPCAIVERERG--WPDFLRCTPDHFPE 180
           :||::: | | :: |    | ::||   :  || |    |  |||  | |  |:||:
Qy      74 FLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLEC--DRFPQ 127
```

RESULT 10
ENTRY S22812 #type complete
TITLE DNA-directed RNA polymerase (EC 2.7.7.6) III largest chain - Giardia lamblia
ORGANISM #formal_name Giardia lamblia
DATE 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 12-Sep-1997
ACCESSIONS S22812
REFERENCE S22811
#authors Lanzendoerfer, M.; Palm, P.; Grampp, B.; Peattie, D.A.; Zillig, W.
#journal Nucleic Acids Res. (1992) 20:1145
#title Nucleotide sequence of the gene encoding the largest subunit of the DNA-dependent RNA polymerase III of Giardia lamblia.
#cross-references MUID:92195823
#accession S22812
##status nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-1735 ##label LAN
##cross-references EMBL:X60325
##note the nucleotide sequence was submitted to the EMBL Data Library, September 1991
GENETICS
#gene rpoA3
CLASSIFICATION #superfamily Giardia DNA-directed RNA polymerase III largest chain
KEYWORDS DNA binding; nucleotidyltransferase; nucleus; transcription; zinc finger
SUMMARY #length 1735 #molecular-weight 193191 #checksum 769

Query Match 4.9%; Score 101; DB 2; Length 1735;
Best Local Similarity 27.4%; Pred. No. 2.94e-01;
Matches 20; Conservative 25; Mismatches 23; Indels 5; Gaps 5;

```
Db     168 TICKKSQGFRV-LREMHYEERATKILADVIKHQKDIEKIQELDQR-LRNHVVDPIQALHI 225
           |:||:  :::: ::|: | :| |||: :  |  | | |:  : :| |:: |:   ::|:
Qy     168 TLCKNDFALKIKVKEITYINRDTKIILET-KS-KTIYKLNGVSERDLKKSVLWLKDSLQC 225

Db     226 L-QKVPECDYPYL 237
             : : : : |||
Qy     226 TCEEMNDINAPYL 238
```

RESULT 11
ENTRY H64825 #type complete
TITLE hypothetical protein b0872 - Escherichia coli (strain K-12)
ORGANISM #formal_name Escherichia coli
DATE 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 14-Nov-1997
ACCESSIONS H64825
REFERENCE A64720
#authors Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.
#journal Science (1997) 277:1453-1462
#title The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97426617
#accession H64825
##status preliminary; nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-322 ##label BLAT
##cross-references GB:AE000189; GB:U00096; NID:g1787097; PID:g1787098; UWGP:b0872
##experimental_source strain K-12, substrain MG1655
SUMMARY #length 322 #molecular-weight 35740 #checksum 5190

Query Match 4.8%; Score 99; DB 2; Length 322;
Best Local Similarity 27.5%; Pred. No. 5.34e-01;
Matches 19; Conservative 22; Mismatches 21; Indels 7; Gaps 7;

```
Db      87 RDVKRGDYLWLSDAMGEFTCDDKAE-DK-FLLLAA-GCGVTPIMSMRRWLAKNRPQADVR 143
           ||:|::  ||| |:: : ||::: : :  :|:::    |   | |::||  | : :   |
Qy     210 RDLKKSV-LWLKDSL-QCTCEEMNDINAPYLVMGQKQGGELVITSVKRW-QKGQREFK-R 265

Db     144 VIYNVRTPQ 152
           :  ::|  |
Qy     266 ISRSIRKLQ 274
```

RESULT 12
ENTRY D31579 #type complete
TITLE cytochrome P450 2D4 - rat
ALTERNATE_NAMES cytochrome P450CMF3
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 28-Feb-1990 #sequence_revision 17-May-1996 #text_change 23-Jan-1998
ACCESSIONS S16873; D31579
REFERENCE S16871
#authors Matsunaga, E.; Umeno, M.; Gonzalez, F.J.
#journal J. Mol. Evol. (1990) 30:155-169
#title The rat P450 IID subfamily: complete sequences of four closely linked genes and evidence that gene conversions maintained sequence homogeneity at the heme-binding region of the cytochrome P450 active site.
#cross-references MUID:90189185
#accession S16873
##molecule_type DNA
##residues 1-500 ##label MAT
##cross-references EMBL:X52029; NID:g57815; PID:g57816
REFERENCE A90151
#authors Ishida, N.; Tawaragi, Y.; Inuzuka, C.; Sugita, O.; Kubota, I.; Nakazato, H.; Noguchi, T.; Sassa, S.
#journal Biochem. Biophys. Res. Commun. (1988) 156:681-688
#title Four species of cDNAs for cytochrome P450 isozymes immunorelated to P450c-M/F encode for members of P450IID subfamily, increasing the number of members within the subfamily.
#cross-references MUID:89050091
#accession D31579
##molecule_type mRNA
##residues 177-500 ##label ISH
##cross-references EMBL:M22331; NID:g203829; PID:g203830
GENETICS
#gene CYP2D4
#introns 63/3; 121/1; 172/1; 225/3; 284/3; 332/1; 394/3; 442/1
CLASSIFICATION #superfamily cytochrome P450
KEYWORDS chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; monooxygenase; oxidoreductase; transmembrane protein
FEATURE
446 #binding_site heme iron (Cys) (axial ligand) #status predicted
SUMMARY #length 500 #molecular-weight 56697 #checksum 7863

Query Match 4.8%; Score 100; DB 2; Length 500;
Best Local Similarity 26.7%; Pred. No. 3.97e-01;
Matches 16; Conservative 22; Mismatches 18; Indels 4; Gaps 3;

```
Db     173 FPFS-PNTLLDKAVCNVIASLLFACR-FEYNDPRFIRLL--DLLKDTLEEESGFLPMLLN 228
           | |: |:    :: |: |:: |  |: :|| : |:  ||  : :|: ||: :: :|::::
Qy       6 FLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMK 65
```

RESULT 13

```
ENTRY            I52313     #type complete
TITLE            cytochrome P450 2D4 homolog, brain (clone 2d-29) - rat
ORGANISM         #formal_name Rattus norvegicus #common_name Norway rat
DATE             27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change
                   23-Jan-1998
ACCESSIONS       I52313
REFERENCE        I52313
   #authors      Kawashima, H.; Strobel, H.W.
   #journal      Biochem. Biophys. Res. Commun. (1995) 209:535-540
   #title        cDNA cloning of a novel rat brain cytochrome P450 belonging
                   to the CYP2D subfamily.
   #cross-references MUID:95251650
   #accession    I52313
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues      1-500 ##label RES
      ##cross-references GB:S77859; NID:g998524; PID:g998525
      ##experimental_source brain, strain Sprague-Dawley
CLASSIFICATION   #superfamily cytochrome P450
KEYWORDS         heme; transmembrane protein
FEATURE
   446                #binding_site heme iron (Cys) (axial ligand) #status
                        predicted
SUMMARY          #length 500  #molecular-weight 56683  #checksum 8540

  Query Match             4.8%;  Score 100;  DB 2;  Length 500;
  Best Local Similarity  26.7%;  Pred. No. 3.97e-01;
  Matches    16;  Conservative   22;  Mismatches  18;  Indels   4;  Gaps   3;

Db    173 FPFS-PNTLLDKAVCNVIASLLFACR-FEYNDPRFIRLL--DLLKDTLEEESGFLPMLLN 228
          | |: |:    :: |: |:: |  |: :|| : |:  ||  : :|: ||: :: :|::::
Qy      6 FLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMK 65


RESULT   14
ENTRY            B32571     #type complete
TITLE            ribosomal protein S6 kinase II (EC 2.7.-.-) alpha chain
                   homolog (clone Mu6A) - mouse
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
                   26-Feb-1998
ACCESSIONS       B32571
REFERENCE        A93113
   #authors      Alcorta, D.A.; Crews, C.M.; Sweet, L.J.; Bankston, L.; Jones,
                   S.W.; Erikson, R.L.
   #journal      Mol. Cell. Biol. (1989) 9:3850-3859
   #title        Sequence and expression of chicken and mouse rsk: homologs of
                   Xenopus laevis ribosomal S6 kinase.
   #cross-references MUID:89384612
   #accession    B32571
      ##molecule_type mRNA
      ##residues      1-724 ##label ALC
      ##cross-references GB:M23489
CLASSIFICATION   #superfamily ribosomal protein S6 kinase II; protein kinase
                   homology
KEYWORDS         ATP; phosphotransferase; serine/threonine-specific protein
                   kinase
FEATURE
   60-310             #domain protein kinase homology #label KIN1\
   68-76              #region protein kinase ATP-binding motif\
   405-664            #domain protein kinase homology #label KIN2
SUMMARY          #length 724  #molecular-weight 81594  #checksum 4206

  Query Match             4.8%;  Score 99;  DB 2;  Length 724;
  Best Local Similarity  35.2%;  Pred. No. 5.34e-01;
  Matches    19;  Conservative   13;  Mismatches  18;  Indels   4;  Gaps   4;

Db    489 KI-LRQKFFSEREASFVLHTISKTVEYLHSQGVVHRDLKPSNILYVDESGNPEC 541
          || :::  : :|:: ::| | |||: |    ||  |||| | :| : :|    |
Qy    177 KIKVKEITYINRDTKIILETKSKTI-Y-KLNGVSERDLKKS-VLWLKDSLQCTC 227


RESULT   15
ENTRY            E64474     #type complete
TITLE            hypothetical protein MJ1398 - Methanococcus jannaschii
ORGANISM         #formal_name Methanococcus jannaschii
DATE             13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
                   10-Oct-1997
ACCESSIONS       E64474
REFERENCE        A64300
   #authors      Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,
                   R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton,
                   R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.;
                   Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;
                   Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
                   Scott, J.L.; Geoghagen, N.S.M.; Weidman, J.F.; Fuhrmann,
                   J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson,
                   J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,
                   K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk,
                   H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
   #journal      Science (1996) 273:1058-1073
   #title        Complete genome sequence of the methanogenic archaeon,
                   Methanococcus jannaschii.
   #cross-references MUID:96337999
   #accession    E64474
      ##status        preliminary; nucleic acid sequence not shown;
                       translation not shown
      ##molecule_type DNA
      ##residues      1-391 ##label BUL
      ##cross-references GB:U67580; GB:L77117; NID:g1592044; PID:g1592046;
                          TIGR:MJ1398; PID:g1511417
GENETICS
   #map_position FOR1358185-1359360
SUMMARY          #length 391  #molecular-weight 45529  #checksum 6984

  Query Match             4.7%;  Score 98;  DB 2;  Length 391;
  Best Local Similarity  30.0%;  Pred. No. 7.17e-01;
  Matches    15;  Conservative   19;  Mismatches  14;  Indels   2;  Gaps   2;

Db     84 LRFKIENKSYNLIITKKNVNTNKMAIIFTSINGSREVNGNKTIIWIKDPL 133
          |::|: : :|    ||  ::|: ::| :  :||  | : :|:::|:||:|
Qy    176 LKIKVKEITYINRDTKIILETKSKTI-YK-LNGVSERDLKKSVLWLKDSL 223


Search completed: Thu Oct 22 15:59:48 1998
Job time : 38 secs.
```

****

MPSRCH (TM)

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Oct 22 15:51:29 1998; MasPar time 18.49 Seconds
794.551 Million cell updates/sec
Tabular output not generated.

Title: >US-08-848-439-2
Description: (1-295) from US08848439.pep
Perfect Score: 2234
Sequence: 1 MLQGPGSLLLLFLASHCCLG.....WQKGQREFKRISRSIRKLQC 295

Scoring table: PAM 150
Gap 11

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembl6
1:sp_archea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 45.109; Variance 71.230; scale 0.633

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	2190	98.0	295	11	O08862	SECRETED FRIZZLED RELA	0.00e+00
2	2188	97.9	295	11	O35297	SECRETED APOPTOSIS REL	0.00e+00
3	2185	97.8	295	11	P97299	STROMAL CELL DERIVED F	0.00e+00
4	1538	68.8	206	4	O14778	SECRETED APOPTOSIS REL	0.00e+00
5	913	40.9	314	4	O14779	SECRETED APOPTOSIS REL	2.63e-199
6	911	40.8	313	4	O00546	SECRETED FRIZZLED-RELA	8.79e-199
7	902	40.4	308	6	O19116	FRZA PRECURSOR.	2.00e-196
8	901	40.3	317	4	O14780	SECRETED APOPTOSIS REL	3.65e-196
9	890	39.8	314	11	O08861	SECRETED FRIZZLED RELA	2.76e-193
10	729	32.6	307	13	O42397	CRESCENT.	2.17e-151
11	582	26.1	281	13	O73821	SECRETED XWNT8 INHIBIT	1.20e-113
12	400	17.9	568	5	O16147	MOM-5.	4.58e-68
13	397	17.8	592	13	O57328	7-TRANSMEMBRANE PROTEI	2.49e-67
14	396	17.7	572	11	Q61090	FRIZZLED HOMOLOG 7 (TR	4.38e-67
15	394	17.6	591	4	O00144	FRIZZLED HOMOLOG.	1.35e-66
16	394	17.6	685	11	Q61091	FRIZZLED HOMOLOG 8 (TR	1.35e-66
17	391	17.5	567	13	O57329	7-TRANSMEMBRANE PROTEI	7.33e-66
18	392	17.5	626	11	O70421	FRIZZLED-1.	4.17e-66
19	386	17.3	565	4	Q14332	FRIZZLED GENE PRODUCT.	1.22e-64
20	386	17.3	570	11	Q08464	FRIZZLED PROTEIN HOMOL	1.22e-64
21	386	17.3	641	11	Q08463	FRIZZLED PROTEIN HOMOL	1.22e-64
22	374	16.7	585	4	Q13467	TRANSMEMBRANE RECEPTOR	1.03e-61
23	369	16.5	537	11	Q61088	FRIZZLED HOMOLOG 4 (TR	1.68e-60
24	357	16.0	525	5	Q10662	FRIZZLED PROTEIN HOMOL	1.36e-57
25	355	15.9	694	5	Q94916	DFZ2.	4.13e-57
26	347	15.5	586	5	Q24760	TISSUE POLARITY PROTEI	3.50e-55
27	337	15.1	558	5	Q94132	TRANSMEMBRANE RECEPTOR	8.84e-53
28	317	14.2	666	11	Q61086	FRIZZLED-3.	5.23e-48
29	309	13.8	664	13	O42579	FRIZZLED 3 PROTEIN.	4.12e-46
30	295	13.2	605	5	O16190	F27E11.3 PROTEIN.	8.19e-43
31	293	13.1	325	6	Q95117	FRZB PRECURSOR.	2.41e-42
32	291	13.0	325	4	O00181	FRITZ.	7.10e-42
33	289	12.9	325	4	Q92765	FRZB PRECURSOR.	2.08e-41
34	289	12.9	325	4	Q99686	FREZZLED.	2.08e-41
35	284	12.7	709	11	Q61089	FRIZZLED HOMOLOG 6 (TR	3.07e-40
36	281	12.6	348	11	O35222	FRIZZLED RELATED PROTE	1.53e-39
37	275	12.3	323	11	P97401	SECRETED FRIZZLED-RELA	3.79e-38
38	274	12.3	346	4	O14877	FRPHE.	6.47e-38
39	267	12.0	318	13	P79936	FREZZLED.	2.68e-36
40	267	12.0	319	13	P79993	FRZB PRECURSOR.	2.68e-36
41	268	12.0	706	4	O60353	FRIZZLED-6.	1.58e-36
42	240	10.7	261	11	O08570	FRITZ (FRAGMENT).	3.87e-30
43	195	8.7	641	4	O00520	CARBOXYPEPTIDASE Z PRE	3.16e-20
44	178	8.0	505	11	Q60569	PROCOLLAGEN, TYPE XVII	1.25e-16
45	178	8.0	562	11	Q60672	PROCOLLAGEN, TYPE XVII	1.25e-16

ALIGNMENTS

RESULT 1
ID O08862 PRELIMINARY; PRT; 295 AA.
AC O08862;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE SECRETED FRIZZLED RELATED PROTEIN SFRP-2.
GN SFRP2.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97250455.
RA RATTNER A., HSIEH J.C., SMALLWOOD P.M., GILBERT D.J., COPELAND N.G.,
RA JENKINS N.A., NATHANS J.;
RL PROC. NATL. ACAD. SCI. U.S.A. 94:2859-2863(1997).
DR EMBL; U88567; G1946343; -.
SQ SEQUENCE 295 AA; 33469 MW; 236B0DDA CRC32;

Query Match 98.0%; Score 2190; DB 11; Length 295;
Best Local Similarity 98.0%; Pred. No. 0.00e+00;
Matches 289; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Db      1 MPRGPASLLLLVLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRL 60
          | :||:||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRL 60

Db     61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQ 120

Db    121 VKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKTKNE 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||:
Qy    121 VKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKND 180

Db    181 DDNDIMETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 DDNDIMETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLK 240

Db    241 DSLQCTCEEMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 DSLQCTCEEMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 295
```

```
RESULT    2
ID   O35297      PRELIMINARY;      PRT;   295 AA.
AC   O35297;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED APOPTOSIS RELATED PROTEIN 1.
GN   SARP1.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M.,
RA   FITZPATRIC P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF017989; G2415421; -.
SQ   SEQUENCE   295 AA;  33483 MW;  0E4C8A8E CRC32;

  Query Match           97.9%;  Score 2188;  DB 11;  Length 295;
  Best Local Similarity  97.6%;  Pred. No. 0.00e+00;
  Matches   288;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db      1 MPRGPASLLLLVLASHCCLGSARGLFLFGQPDFSYKRTNCKPIPANLQLCHGIEYQNMRL 60
          | :||:|||||  ||||||||||||||||||||||:||||||||||||||||||||||||
Qy      1 MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRL 60

Db     61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQ 120

Db    121 VKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKTKNE 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||:
Qy    121 VKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKND 180

Db    181 DDNDIMETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 DDNDIMETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLK 240

Db    241 DSLQCTCEEMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 DSLQCTCEEMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 295


RESULT    3
ID   P97299      PRELIMINARY;      PRT;   295 AA.
AC   P97299;
DT   01-MAY-1997 (TREMBLREL. 03, CREATED)
DT   01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   STROMAL CELL DERIVED FACTOR 5 (SDF5).
GN   SDF5.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97092876.
RA   SHIROZU M., TADA H., TASHIRO K., NAKAMURA T., LOPEZ N.D., NAZAREA M.,
RA   HAMADA T., SATO T., NAKANO T., HONJO T.;
RL   GENOMICS 37:273-280(1996).
DR   EMBL; D50462; G1747302; -.
DR   MGD; MGI:108078; SDF5.
SQ   SEQUENCE   295 AA;  33501 MW;  329888C8 CRC32;

  Query Match           97.8%;  Score 2185;  DB 11;  Length 295;
  Best Local Similarity  97.6%;  Pred. No. 0.00e+00;
  Matches   288;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db      1 MPRGPASLLLLVLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRL 60
          | :||:|||||  ||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRL 60

Db     61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSRCVQ 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy     61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQ 120

Db    121 VKDRCAPVMSAF--PWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKND 178
          |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 VKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKND 180

Db    179 DDNDIMETLCKNDFALKIKVKEITYINR 206
          |||||||||||||||||||||||||||||
Qy    181 DDNDIMETLCKNDFALKIKVKEITYINR 208


RESULT    5
ID   O14779      PRELIMINARY;      PRT;   314 AA.
AC   O14779;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED APOPTOSIS RELATED PROTEIN 2.
GN   SARP2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
```

```
Db     61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCMQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Qy     61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQ 120

Db    121 VKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKTKNE 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||:
Qy    121 VKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKND 180

Db    181 DDNDIMETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 DDNDIMETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLK 240

Db    241 DSLQCTCEEMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 DSLQCTCEEMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC 295


RESULT    4
ID   O14778      PRELIMINARY;      PRT;   206 AA.
AC   O14778;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED APOPTOSIS RELATED PROTEIN 1 (FRAGMENT).
GN   SARP1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M.,
RA   FITZPATRIC P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF017986; G2415415; -.
FT   NON_TER     206     206
SQ   SEQUENCE   206 AA;  23215 MW;  F465B68B CRC32;

  Query Match           68.8%;  Score 1538;  DB 4;  Length 206;
  Best Local Similarity  98.6%;  Pred. No. 0.00e+00;
  Matches   205;  Conservative    0;  Mismatches   1;  Indels   2;  Gaps   1;

Db      1 MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRL 60

Db     61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSRCVQ 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy     61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQ 120

Db    121 VKDRCAPVMSAF--PWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKND 178
          |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 VKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKND 180

Db    179 DDNDIMETLCKNDFALKIKVKEITYINR 206
          |||||||||||||||||||||||||||||
Qy    181 DDNDIMETLCKNDFALKIKVKEITYINR 208
```

```
RA   MELKONYAN H., PROCHAZKA V., CHANG W.C., UMANSKY S.R.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF017987; G2415417; -.
SQ   SEQUENCE   314 AA;  35411 MW;  61E261B5 CRC32;

  Query Match           40.9%;  Score 913;  DB 4;  Length 314;
  Best Local Similarity 42.6%;  Pred. No. 2.63e-199;
  Matches  112;  Conservative   67;  Mismatches  77;  Indels   7;  Gaps   5;

Db      51 FYTKPPQCVDIPADLRLCHNVGYKKMVLPNLLEHETMAEVKQQASSWVPLLNKNCHAGTQ 110
           |  | : |  |||:|:||| : | :| ||||| |||| || :||::|:||: | ||: |
Qy      33 FSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 92

Db     111 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGFYWPEMLKCDKFP-EGDVC 166
            ||||||||||||     | ||: ||  |:| | |||  ||| ||:|| ||:|| : |:|
Qy      93 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 152

Db     167 IAMTPPNPTEASKPQGTTVCPPCDNELKSEA-IIEHLCASEFALRMKIKEVKKENGDKKI 225
           |::::::   ::  ::  || :| |    :  |:| || ::|||::|:||:   | | ||
Qy     153 IPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKI 212

Db     226 V-PKKKKPL-KLGPIKKKDLKKLVLYLKNGADCPCHQLDNLSHHFLIMGRKVKSQYLLTA 283
           :   | | : ||  :  :|||| || ||:: :| | :::::: :|:||:|  :: ::|:
Qy     213 ILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 272

Db     284 IHKWDKKNKEFKNFMKKMKNHEC 306
           : :|:|  :||| : : ::: :|
Qy     273 VKRWQKGQREFKRISRSIRKLQC 295


RESULT    6
ID   O00546      PRELIMINARY;      PRT;   313 AA.
AC   O00546;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED FRIZZLED-RELATED PROTEIN.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   FINCH P.W., HE X., KELLEY M.J., UREN A., SCHAUDIES R.P., POPESCU N.C.,
RA   RUDIKOFF S., AARONSON S.A., VARMUS H.E., RUBIN J.S.;
RL   SUBMITTED (JUN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   ZHOU Z., WANG J.;
RL   SUBMITTED (MAR-1998) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF001900; G2213819; -.
DR   EMBL; AF056087; G3033551; -.
SQ   SEQUENCE   313 AA;  35314 MW;  A3CD04FF CRC32;

  Query Match           40.8%;  Score 911;  DB 4;  Length 313;
  Best Local Similarity 42.6%;  Pred. No. 8.79e-199;
  Matches  112;  Conservative   67;  Mismatches  77;  Indels   7;  Gaps   5;

Db      50 FYTKPPQCVDIPADLRLCHNVGYKKMVLPNLLEHETMAEVKQQASSWVPLLNKNCHAGTQ 109
           |  | : |  |||:|:||| : | :| ||||| |||| || :||::|:||: | ||: |
Qy      33 FSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 92

Db     110 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGFYWPEMLKCDKFP-EGDVC 165
            ||||||||||||     | ||: ||  |:| | |||  ||| ||:|| ||:|| : |:|
Qy      93 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 152

Db     166 IAMTPPNATEASKPQGTTVCPPCDNELKSEA-IIEHLCASEFALRMKIKEVKKENGDKKI 224
           |::::::   ::  ::  || :| |    :  |:| || ::|||::|:||:   | | ||
Qy     153 IPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKI 212

Db     225 V-PKKKKPL-KLGPIKKKDLKKLVLYLKNGADCPCHQLDNLSHHFLIMGRKVKSQYLLTA 282
           :   | | : ||  :  :|||| || ||:: :| | :::::: :|:||:|  :: ::|:
Qy     213 ILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 272

Db     283 IHKWDKKNKEFKNFMKKMKNHEC 305
           : :|:|  :||| : : ::: :|
Qy     273 VKRWQKGQREFKRISRSIRKLQC 295


RESULT    7
ID   O19116      PRELIMINARY;      PRT;   308 AA.
AC   O19116;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   FRZA PRECURSOR.
OS   BOS TAURUS (BOVINE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RA   DUPLAA C., D'AMORE P.A.;
RL   SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U85945; G2337937; -.
KW   SIGNAL.
FT   SIGNAL        5     22       POTENTIAL.
SQ   SEQUENCE   308 AA;  34763 MW;  9B74A92F CRC32;

  Query Match           40.4%;  Score 902;  DB 6;  Length 308;
  Best Local Similarity 42.2%;  Pred. No. 2.00e-196;
  Matches  111;  Conservative   66;  Mismatches  79;  Indels   7;  Gaps   5;

Db      45 FYTKPPQCVDIPADLRLCHNVGYKRMVLPNLLEHETMAEVKQQASSWVPLLNKNCHIGTQ 104
           |  | : |  |||:|:||| : |  | ||||| |||| || :||::|:||: | ||  |
Qy      33 FSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 92

Db     105 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGFYWPEMLKCDKFP-EGDVC 160
            ||||||||||||     | ||: ||  |:| | |||  ||| ||:|| ||:|| : |:|
Qy      93 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 152

Db     161 IAMTPPNATEASKPQGTTVCPPCDNELKSEA-IIEHLCASEFALRMKIKEVKKENGDKKI 219
           |::::::   ::  ::  || :| |    :  |:| || ::|||::|:||:   | | ||
Qy     153 IPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKI 212

Db     220 V-PKKKKPL-KLGPIKKKELKKLVLYLKNGADCPCHQLDNLSHHFLIMGRKVKSQYLLTA 277
           :   | | : ||  :  ::||| || ||:: :| | :::::: :|:||:|  :: ::|:
Qy     213 ILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 272

Db     278 IHKWDKKNKEFKTFMKKMKNHEC 300
           : :|:|  :||| : : ::: :|
Qy     273 VKRWQKGQREFKRISRSIRKLQC 295


RESULT    8
ID   O14780      PRELIMINARY;      PRT;   317 AA.
AC   O14780;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED APOPTOSIS RELATED PROTEIN 3.
GN   SARP3.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PANCREAS;
RA   MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M.,
RA   FITZPATRIC P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF017988; G2415419; -.
SQ   SEQUENCE   317 AA;  35577 MW;  2CD69373 CRC32;

  Query Match           40.3%;  Score 901;  DB 4;  Length 317;
  Best Local Similarity 45.2%;  Pred. No. 3.65e-196;
```

```
  Matches   114;  Conservative   58;  Mismatches  74;  Indels   6;  Gaps   4;

Db      46 SYSKPPQCLDIPADLPLCHTVGYKRMRLPNLLEHESLAEVKQQASSWLPLLAKRCHSDTQ 105
           || | : |  |||:| ||| : |  ||||||| ||:: || :||::|:||: |:||:||
Qy      34 SY-KRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 92

Db     106 VFLCSLFAPVCLD---RPIYPCRSLCEAVRAGCAPLMEAYGFPWPEMLHCHKFPLDNDLC 162
            ||||||||||||     | ||:|||  |:  |||:| |:|||||:|| | :|| |||||
Qy      93 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 152

Db     163 IAVQFG-HLPATAPPVTKICAQC-EMEHSADGLMEQMCSSDFVVKMRIKEIKIENGDRKL 220
           |::  : || :::    |:|  | :::   : :|| :| :|| :|:::|||   | | |:
Qy     153 IPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKI 212

Db     221 IGAQKKKKLLKPGPLKRKDTKRLVLHMKNGAGCPCPQLDSLAGSFLVMGRKVDGQLLLMA 280
           |   | | : |   :  :| |: || :|:: | | ::: : :::|||||:|  |:|:: :
Qy     213 ILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 272

Db     281 VYRWDKKNKEMK 292
           | ||:|  :| |
Qy     273 VKRWQKGQREFK 284


RESULT    9
ID   O08861      PRELIMINARY;      PRT;   314 AA.
AC   O08861;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   SECRETED FRIZZLED RELATED PROTEIN SFRP-1.
GN   SFRP1.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97250455.
RA   RATTNER A., HSIEH J.C., SMALLWOOD P.M., GILBERT D.J., COPELAND N.G.,
RA   JENKINS N.A., NATHANS J.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 94:2859-2863(1997).
DR   EMBL; U88566; G1946341; -.
SQ   SEQUENCE   314 AA;  35327 MW;  5AA5CA27 CRC32;

  Query Match           39.8%;  Score 890;  DB 11;  Length 314;
  Best Local Similarity 41.4%;  Pred. No. 2.76e-193;
  Matches   109;  Conservative   67;  Mismatches  80;  Indels   7;  Gaps   5;

Db      51 FYTKPPQCVDIPVDLRLCHNVGYKKMVLPNLLEHETMAEVKQQASSWVPLLNKNCHMGTQ 110
           |  | : |  || :|:||| : | :| ||||| |||| || :||::|:||: | ||  |
Qy      33 FSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK 92

Db     111 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGFYWPEMLKCDKFP-EGDVC 166
            ||||||||||||     | ||: ||  |:| | |||  ||| ||:|| ||:|| : |:|
Qy      93 KFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLC 152

Db     167 IAMTPPNTTEASKPQGTTVCPPCDNELKSEA-IIEHLCASEFALRMKIKEVKKENGDKKI 225
           |::::::   ::  ::  || :| |    :  |:| || ::|||::|:||:   | | ||
Qy     153 IPLASSDHLLPATEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTKI 212

Db     226 V-PKKKKPL-KLGPIKKKELKALVLFLKNGADCPCHQLDNLSHNFLIMGRKVKSQYLLTA 283
           :   | | : ||  :  ::|| || ||:: :| | :::::  :|:||:|  :: ::|:
Qy     213 ILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 272

Db     284 IHKWDKKNKEFKNFMKRMKNHEC 306
           : :|:|  :||| : : ::: :|
Qy     273 VKRWQKGQREFKRISRSIRKLQC 295


RESULT   10
ID   O42397      PRELIMINARY;      PRT;   307 AA.
AC   O42397;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   CRESCENT.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   PFEFFER P.L., IZPISUA-BELMONTE J.C., DE ROBERTIS E.M.;
RL   SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF006508; G2226372; -.
SQ   SEQUENCE   307 AA;  34683 MW;  320A2243 CRC32;

  Query Match           32.6%;  Score 729;  DB 13;  Length 307;
  Best Local Similarity 35.8%;  Pred. No. 2.17e-151;
  Matches    96;  Conservative   68;  Mismatches  94;  Indels  10;  Gaps   8;

Db      36 SYLRRSSSCTAIPRSMALCYDIGYSEMRIPNLLEHETMPEVIQQSSSWLPLLARECHPDA 95
           || :||: | :|| :: ||  | | :||:|||| |||| ||::|:::|:||: ::|||||:
Qy      34 SY-KRSN-CKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDT 91

Db      96 RIFLCSLFAPICLDRL---IYPCRSLCEAVKRSCAPVMACYGYPWPEILNCNKFPADHEL 152
           : |||||||||:||| |   | ||:|||  ||  |||||: :|:|||::|:|::||  |::|
Qy      92 KKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDL 151

Db     153 CIAAVSTDES-SSSRRMPRASCKDCELEEASTAREILDNLCANDFTVKIRILRKNTTTTV 211
           ||:  |:|   :::   |:  |  |  :      :|:: || |||::||::
Qy     152 CIPLASSDHLLPATEEAPKV-CEACK-NKNDDDNDIMETLCKNDFALKIKVKEITYINRD 209

Db     212 SDFDLDPSRVEVLKPGPLLRTEIPGRLQQWLDIDATCAHNIMRGTHAGVFVISGEVRSDK 271
           : : |:     : |   :   ::   :  ||     |: : |   :|  ::: |: ::
Qy     210 TKIILETKSKTIYKLNGVSERDLKKSVL-WLKDSLQCTCEEMNDINAP-YLVMGQKQGGE 267

Db     272 VVVNKAYAWQKKNRNLHQAVRRWKHHRC 299
           :|:     |||  |:: :  |  :  :|
Qy     268 LVITSVKRWQKGQREFKRISRSIRKLQC 295


RESULT   11
ID   O73821      PRELIMINARY;      PRT;   281 AA.
AC   O73821;
DT   01-AUG-1998 (TREMBLREL. 07, CREATED)
DT   01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   SECRETED XWNT8 INHIBITOR SIZZLED.
GN   SZL.
OS   XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 98088686.
RA   SALIC A.N., KROLL K.L., EVANS L.M., KIRSCHNER M.W.;
RL   DEVELOPMENT 124:4739-4748(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   SALIC A.N., KROLL K.L., EVANS L.M., KIRSCHNER M.W.;
RL   SUBMITTED (APR-1998) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF059570; G3089551; -.
SQ   SEQUENCE   281 AA;  31834 MW;  8478D86B CRC32;

  Query Match           26.1%;  Score 582;  DB 13;  Length 281;
  Best Local Similarity 32.7%;  Pred. No. 1.20e-113;
  Matches    86;  Conservative   65;  Mismatches 101;  Indels  11;  Gaps   8;

Db      25 TKCVTIPTEMAMCNDVGYSEMRLPNLMGHTNMAEVVPKSAEWQNLLQTGCHPYARTFLCS 84
           ::|  ||::: :|: : | :|||||:||  | ||:  :: |  |:   ||| :: ||||
Qy      38 SNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCS 97

Db      85 LFAPVCLDTF---IQPCRSMCVAVRDSCAPVLACHGHSWPESLDCDRFPAGEDMCLDTLS 141
           |||||||| :   ||||:|:|| |:| ||||::  | :||: |:|||||  :|:|:   |
Qy      98 LFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLAS 157

Db     142 KEYQY-SYKELPKPSCQGCPLIEEFFSHKTVLEAFCDNNFAVKVKLAKKKSAS-GL-YEY 198
           :    :  | ||  |::|   ::    : ::|::| |:||:|:|:      :
```

```
Qy    158 SDHLLPATEEAPK-VCEACKNKND--DDNDIMETLCKNDFALKIKVKEITYINRDTKIIL 214

Db    199 ETEGPVEFIKQGLLLPYDTRTMIEQWLLINENCAQKLIRTRPTVYVIAGEIHHGKVKVNR 258
          || :   :   |:    | :  :  ||  :  |:   :    : |:: |: : | : :
Qy    215 ETKSKTIYKLNGVSER-DLKKSV-LWLKDSLQCTCEEMNDINAPYLVMGQKQGGELVITS 272

Db    259 IFHWQKKDSQLTLATRRWRHHKC 281
          : :||| : ::   :|  |   |
Qy    273 VKRWQKGQREFKRISRSIRKLQC 295


RESULT   12
ID   O16147      PRELIMINARY;      PRT;   568 AA.
AC   O16147;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   MOM-5.
GN   MOM-5.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL;
RX   MEDLINE; 97433081.
RA   ROCHELEAU C.E., DOWNS W.D., LIN R., WITTMANN C., BEI Y., CHA Y.-H.,
RA   ALI M., PRIESS J.R., MELLO C.C.;
RL   CELL 90:707-716(1997).
DR   EMBL; AF013953; G2463674; -.
SQ   SEQUENCE   568 AA;  62892 MW;  8578BB4F CRC32;

  Query Match           17.9%;  Score 400;  DB 5;  Length 568;
  Best Local Similarity  37.4%;  Pred. No. 4.58e-68;
  Matches    58;  Conservative   33;  Mismatches  54;  Indels  10;  Gaps   8;

Db      5 ILILFLFG--CLSADQRLSSTSISSMNGFSTTR-KCEHI-T-IPMCKNLDYNQTVFPNLL 59
          :|:||| :  ||::  |  :  : :   ||  | :|  | : : :|  ::|    :||||
Qy      8 LLLLFLASHCCLGSA-R--GLFLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLL 64

Db     60 GHTTQSEAGPAIAQFNPLIKVKCSEDIRLFLCTVYAPVC-TVLEKPIQPCRELCLSAKNG 118
          || |  |      :   ||::  |  | : |||:::|||   |:  ||||: ||:  |:
Qy     65 GHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDR 124

Db    119 CESLMKKFGFQWPDQLDCNKFPV-TDLCVGKNSSE 152
          | ::|  ||| ||| |:|::||   |||:   ||:
Qy    125 CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSD 159


RESULT   13
ID   O57328      PRELIMINARY;      PRT;   592 AA.
AC   O57328;
DT   01-JUN-1998 (TREMBLREL. 06, CREATED)
DT   01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   7-TRANSMEMBRANE PROTEIN FRIZZLED-1.
GN   FZ-1.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIMB BUD;
RA   KENGAKU M., TWOMBLY V., TABIN C.;
RL   COLD SPRING HARB. SYMP. QUANT. BIOL. 0:0-0(1997).
DR   EMBL; AF031830; G2655274; -.
KW   TRANSMEMBRANE.
SQ   SEQUENCE   592 AA;  65490 MW;  CD66EC83 CRC32;

  Query Match           17.8%;  Score 397;  DB 13;  Length 592;
  Best Local Similarity  37.7%;  Pred. No. 2.49e-67;
  Matches    49;  Conservative   32;  Mismatches  44;  Indels   5;  Gaps   5;

Db     67 HGYCQPI-S-IPLCTDIAYNQTIMPNLLGHTNQEDAGLEVHQFYPLVKVQCSAELKFFLC 124
          :: | || : : ||  | |    :||||||    :   :     |||: || :: | |||
Qy     37 RSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 96

Db    125 SMYAPVC-TVLEQALPPCRSLCERARQGCEALMNKFGFQWPDTLRCEKFPVHGAGELCVG 183
          |::||||   |:::: ||:||| : :: | ::|: ||| ||| | |::|| :   :||:
Qy     97 SLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QDN-DLCIP 154

Db    184 QNASERGTPT 193
            :|::  |:
Qy    155 LASSDHLLPA 164


RESULT   14
ID   Q61090      PRELIMINARY;      PRT;   572 AA.
AC   Q61090;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   FRIZZLED HOMOLOG 7 (TRANSMEMBRANE RECEPTOR).
GN   FZD7 OR FRIZZLED 7.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96224032.
RA   WANG Y., MACKE J.P., ABELLA B.S., ANDREASSON K., WORLEY P.,
RA   GILBERT D.J., COPELAND N.G., JENKINS N.A., NATHANS J.;
RL   J. BIOL. CHEM. 271:4468-4476(1996).
DR   EMBL; U43320; G1151258; -.
DR   MGD; MGI:108570; FZD7.
KW   TRANSMEMBRANE.
SQ   SEQUENCE   572 AA;  63816 MW;  FA1D78D2 CRC32;

  Query Match           17.7%;  Score 396;  DB 11;  Length 572;
  Best Local Similarity  40.7%;  Pred. No. 4.38e-67;
  Matches    48;  Conservative   25;  Mismatches  42;  Indels   3;  Gaps   3;

Db     51 PISIPLCTDIAYNQTILPNLLGHTNQEDAGLEVHQFYPLVKVQCSPELRFFLCSMYAPVC 110
          | :: ||  | |    |||||||    :   :     |||: || |: : ||||::||||
Qy     44 PANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVC 103

Db    111 -TVLDQAIPPCRSLCERARQGCEALMNKFGFQWPERLRCENFPVHGAGEICVGQNTSD 167
             ||::| ||:||| : :: | ::|: ||| ||: | |: || :   ::|:   :||
Qy    104 LDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QDN-DLCIPLASSD 159


RESULT   15
ID   O00144      PRELIMINARY;      PRT;   591 AA.
AC   O00144;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   FRIZZLED HOMOLOG.
GN   FZD3.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97227293.
RA   WANG Y.K., HARRYMAN SAMOS C., PEOPLES R., PEREZ-JURADO L.A., NUSSE R.,
RA   FRANCKE U.;
RL   HUM. MOL. GENET. 6:465-472(1997).
DR   EMBL; U82169; G1906598; -.
SQ   SEQUENCE   591 AA;  64466 MW;  88C5EC1F CRC32;

  Query Match           17.6%;  Score 394;  DB 4;  Length 591;
  Best Local Similarity  37.5%;  Pred. No. 1.35e-66;
  Matches    48;  Conservative   32;  Mismatches  47;  Indels   1;  Gaps   1;

Db     41 AVEIPMCRGIGYNLTRMPNLLGHTSQGEAAAELAEFAPLVQYGCHSHLRFFLCSLYAPMC 100
          : :: :|:|| |   |:|||||| :  |   : :   |||   ||:  : |||||:||:|
```

```
Qy      44 PANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVC 103

Db     101 TDQVSTPIPACRPMCEQARLRCAPIMEQFNFGWPDSLDCARLPTRNDPHALCMEAPENAT 160
           |::   | :|:::| | : ||||:|  | | ||| |:| |:|  ||   : : ::::
Qy     104 LDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDL-CIPLASSDHLL 162

Db     161 AGPAEPHK 168
           ::  |: |
Qy     163 PATEEAPK 170


Search completed: Thu Oct 22 15:52:56 1998
Job time : 87 secs.
```

****

MPSRCH (TM)

****

Release 3.1A John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Oct 22 16:00:06 1998; MasPar time 8.77 Seconds
507.408 Million cell updates/sec
Tabular output not generated.

Title: >US-08-848-439-3
Description: (1-275) from US08848439.pep
Perfect Score: 2073
Sequence: 1 SARGLFLFGQPDFSYKRSNC.....WQKGQREFKRISRSIRKLQC 275

Scoring table: PAM 150
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 33.049; Variance 130.530; scale 0.253

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	396	19.1	572	27	W31273	Mouse frizzled-7 prot	2.11e-28
2	394	19.0	685	27	W31274	Mouse frizzled-8 prot	3.38e-28
3	374	18.0	585	27	W31271	Human frizzled-5 prot	3.66e-26
4	369	17.8	537	27	W31270	Mouse frizzled-4 prot	1.18e-25
5	357	17.2	525	27	W31269	Caenorhabditis frizzl	1.93e-24
6	355	17.1	694	27	W31267	Drosophila frizzled-2	3.07e-24
7	317	15.3	666	27	W31268	Mouse frizzled-3 prot	2.03e-20
8	284	13.7	709	27	W31272	Mouse frizzled-6 prot	3.89e-17
9	148	7.1	581	14	R74187	Chick p75.	3.20e-04
10	124	6.0	605	14	R74186	Chick p78.	3.95e-02
11	121	5.8	604	28	W35946	Human netrin-1.	7.10e-02
12	105	5.1	529	14	R74188	Mouse p78.	1.50e+00
13	100	4.8	732	25	W26642	Human RECK cancer-inh	3.78e+00
14	99	4.8	804	29	W55368	H. pylori ORF 09ap114	4.53e+00
15	97	4.7	15281	9	R44929	T. niveum Cyclosporin	6.52e+00
16	96	4.6	868	25	W26610	Rat muscle-specific k	7.81e+00
17	96	4.6	868	25	W26507	Rat Dmk receptor.	7.81e+00
18	95	4.6	869	25	W26611	Human muscle-specific	9.35e+00
19	95	4.6	869	25	W26506	Human Dmk receptor.	9.35e+00
20	94	4.5	467	17	R84094	Nsk2 receptor with pu	1.12e+01
21	93	4.5	478	16	R92718	Mouse muscle-localize	1.34e+01
22	93	4.5	530	16	R92715	Mouse muscle-localise	1.34e+01
23	93	4.5	860	16	R92716	Mouse muscle-localise	1.34e+01
24	93	4.5	868	16	R92717	Mouse muscle-localise	1.34e+01
25	87	4.2	343	6	R28887	pRi8 fusion protein.	3.84e+01
26	88	4.2	475	17	R94982	Nsk2 extracellular do	3.23e+01
27	87	4.2	530	6	R28888	Ri paraneoplastic ant	3.84e+01
28	87	4.2	550	13	R72830	Human TIABP2.	3.84e+01
29	88	4.2	580	29	W46758	Amino acid sequence o	3.23e+01
30	88	4.2	580	24	W25670	hNET sequence encoded	3.23e+01
31	88	4.2	740	19	R95911	Insulin-stimulated pr	3.23e+01
32	88	4.2	740	19	R95912	Mutant insulin-stimul	3.23e+01
33	88	4.2	863	17	R84088	Nsk2 receptor with al	3.23e+01
34	88	4.2	871	17	R84087	Nsk2 receptor.	3.23e+01
35	88	4.2	873	17	R84092	Nsk2 receptor with al	3.23e+01
36	88	4.2	881	17	R84091	Nsk2 receptor with al	3.23e+01
37	88	4.2	2516	20	W01875	Neuronal invertebrate	3.23e+01
38	88	4.2	2516	20	W01884	Invertebrate calcium	3.23e+01
39	87	4.2	2695	20	W07632	Human type I inositol	3.84e+01
40	87	4.2	2713	19	W00168	Human T cell inositol	3.84e+01
41	87	4.2	2749	3	R13887	Inositol-3-phosphate	3.84e+01
42	84	4.1	461	7	R37270	ODC.	6.44e+01
43	83	4.0	202	16	R77524	NK-Lysine full length	7.64e+01
44	82	4.0	311	8	R43683	Lipopolysaccharide in	9.04e+01
45	83	4.0	783	24	W27113	Rat spleen pro-hormon	7.64e+01

ALIGNMENTS

RESULT 1
ID W31273 standard; Protein; 572 AA.
AC W31273;
DT 27-APR-1998 (first entry)
DE Mouse frizzled-7 protein Mfz7 (Wnt receptor).
KW Wnt receptor; mouse frizzled-7 protein; Mfz7 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Mus musculus.
PN WO9739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV JOHNS HOPKINS.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR WPI; 97-526631/48.
DR N-PSDB; T89891.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 44-46; 61pp; English.
CC This protein comprises the mouse transmembrane receptor,
CC frizzled-7 (Mfz7), encoded by the Mfz7 gene (see T89891). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 572 AA;

Query Match 19.1%; Score 396; DB 27; Length 572;
Best Local Similarity 40.7%; Pred. No. 2.11e-28;
Matches 48; Conservative 25; Mismatches 42; Indels 3; Gaps 3;

Db 51 pisiplctdiaynqtilpnllghtnqedaglevhqfyplvkvqcspelrfflcsmyapvc 110

```
Qy      24 PANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVC 83

Db     111 -tvldqaippcrslcerarqgcealmnkfgfqwperlrcenfpvhgageicvgqntsd 167
             ||::| ||:||| : :: | ::|: ||| ||: | |: || :   ::|:   :||
Qy      84 LDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QDN-DLCIPLASSD 139


RESULT    2
ID   W31274 standard; Protein; 685 AA.
AC   W31274;
DT   27-APR-1998  (first entry)
DE   Mouse frizzled-8 protein Mfz8 (Wnt receptor).
KW   Wnt receptor; mouse frizzled-8 protein; Mfz8 gene;
KW   signal transduction; cancer; cell growth; cell proliferation.
OS   Mus musculus.
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89892.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Disclosure; Page 48-50; 61pp; English.
CC   This protein comprises the mouse transmembrane receptor,
CC   frizzled-8 (Mfz8), encoded by the Mfz8 gene (see T89892).  It is
CC   an example of a Wnt receptor.  Other novel frizzled family members
CC   have been identified in human, mouse and Caenorhabditis elegans
CC   (see W31268-74) and are considered also to be Wnt receptors.  Wnt
CC   receptors can be used in a novel, claimed method of screening for
CC   compounds which modulate the binding of a Wnt polypeptide (secreted
CC   proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC   Wnt is involved in (mammary) cancer and other processes involving
CC   growth, development and proliferation (both normal and abnormal).
CC   Modulators identified by the claimed method are useful for
CC   treatment of diseases related to these conditions.
SQ   Sequence   685 AA;

  Query Match           19.0%;  Score 394;  DB 27;  Length 685;
  Best Local Similarity  43.8%;  Pred. No. 3.38e-28;
  Matches    56;  Conservative   23;  Mismatches  43;  Indels   6;  Gaps   6;

Db      40 vplckgigynytympnqfnhdtqdeaglevhqfw-plveiqcspdlkfflcsmytpicle 98
           : || || |    :|| : |:|  |  ||    | |||  || || | ||||:::|:||:
Qy      27 LQLCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLD 85

Db      99 dykkplppcrsvcerakagcaplmrqygfawpdrmrcdrlpeqgnpdtlcmdynrtd-lt 157
           |    : ||:|:| : |  |||:|  :||:||| : |||:| | | | ||:    :| |
Qy      86 DLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QDN-D-LCIPLASSDHLL 142

Db     158 taapsppr 165
             |:  :|:
Qy     143 PATEEAPK 150


RESULT    3
ID   W31271 standard; Protein; 585 AA.
AC   W31271;
DT   27-APR-1998  (first entry)
DE   Human frizzled-5 protein Mfz5 (Wnt receptor).
KW   Wnt receptor; human frizzled-5 protein; Hfz5 gene;
KW   signal transduction; cancer; cell growth; cell proliferation.
OS   Homo sapiens.
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89889.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Disclosure; Page 37-39; 61pp; English.
CC   This protein comprises the human transmembrane receptor,
CC   frizzled-5 (Hfz5), encoded by the Hfz5 gene (see T89889).  It is
CC   an example of a Wnt receptor.  Other novel frizzled family members
CC   have been identified in human, mouse and Caenorhabditis elegans
CC   (see W31268-74) and are considered also to be Wnt receptors.  Wnt
CC   receptors can be used in a novel, claimed method of screening for
CC   compounds which modulate the binding of a Wnt polypeptide (secreted
CC   proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC   Wnt is involved in (mammary) cancer and other processes involving
CC   growth, development and proliferation (both normal and abnormal).
CC   Modulators identified by the claimed method are useful for
CC   treatment of diseases related to these conditions.
SQ   Sequence   585 AA;

  Query Match           18.0%;  Score 374;  DB 27;  Length 585;
  Best Local Similarity  38.8%;  Pred. No. 3.66e-26;
  Matches    47;  Conservative   29;  Mismatches  41;  Indels   4;  Gaps   4;

Db      20 lvgraaaaskapvcqei-t-vpmcrgigynlthmpnqfnhdtqdeaglevhqfw-plvei 76
           | |::  : | : |  | : : :|:|| |   ::|| : |:|  |  ||    | |||
Qy       7 LFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMK 65

Db      77 qcspdlrfflctmytpiclpdyhkplppcrsvcerakagcsplmrqygfawpermscdrl 136
           || || : |||::::|:|| |    : ||:|:| : |  |:|:|  :||:||: : |||:
Qy      66 QCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRF 125

Db     137 p 137
           |
Qy     126 P 126


RESULT    4
ID   W31270 standard; Protein; 537 AA.
AC   W31270;
DT   27-APR-1998  (first entry)
DE   Mouse frizzled-4 protein Mfz4 (Wnt receptor).
KW   Wnt receptor; mouse frizzled-4 protein; Mfz4 gene;
KW   signal transduction; cancer; cell growth; cell proliferation.
OS   Mus musculus.
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89888.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Disclosure; Page 34-35; 61pp; English.
CC   This protein comprises the mouse transmembrane receptor,
CC   frizzled-4 (Mfz4), encoded by the Mfz4 gene (see T89888).  It is
CC   an example of a Wnt receptor.  Other novel frizzled family members
CC   have been identified in human, mouse and Caenorhabditis elegans
CC   (see W31268-74) and are considered also to be Wnt receptors.  Wnt
CC   receptors can be used in a novel, claimed method of screening for
CC   compounds which modulate the binding of a Wnt polypeptide (secreted
CC   proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC   Wnt is involved in (mammary) cancer and other processes involving
CC   growth, development and proliferation (both normal and abnormal).
CC   Modulators identified by the claimed method are useful for
```

```
CC   treatment of diseases related to these conditions.
SQ   Sequence   537 AA;

  Query Match           17.8%;  Score 369;  DB 27;  Length 537;
  Best Local Similarity  38.3%;  Pred. No. 1.18e-25;
  Matches    41;  Conservative   26;  Mismatches  40;  Indels   0;  Gaps   0;

Db      47 piriamcqnlgynvtkmpnlvghelqtdaelqlttftpliqygcssqlqfflcsvyvpmc 106
           |  : :|: : |   ::|||:|||   :   |  :  ||:   | ::   ||||:: |:|
Qy      24 PANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVC 83

Db     107 tekinipigpcggmclsvkrrcepvlrefgfawpdtlncskfppqnd 153
            : ::  | || ::|: || || ||:  |||:||| |:| :|| :||
Qy      84 LDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDND 130


RESULT    5
ID   W31269 standard; Protein; 525 AA.
AC   W31269;
DT   27-APR-1998  (first entry)
DE   Caenorhabditis frizzled-1 protein Cfz1 (Wnt receptor).
KW   Wnt receptor; Caenorhabditis frizzled-1 protein; Cfz1 gene;
KW   signal transduction; cancer; cell growth; cell proliferation.
OS   Caenorhabditis elegans.
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89887.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Disclosure; Page 30-31; 61pp; English.
CC   This protein comprises the Caenorhabditis putative transmembrane
CC   receptor, frizzled-1, encoded by the Cfz1 gene (see T89887).  It is
CC   an example of a Wnt receptor.  Other novel frizzled family members
CC   have been identified in human, mouse and Caenorhabditis elegans
CC   (see W31268-74) and are considered also to be Wnt receptors.  Wnt
CC   receptors can be used in a novel, claimed method of screening for
CC   compounds which modulate the binding of a Wnt polypeptide (secreted
CC   proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC   Wnt is involved in (mammary) cancer and other processes involving
CC   growth, development and proliferation (both normal and abnormal).
CC   Modulators identified by the claimed method are useful for
CC   treatment of diseases related to these conditions.
SQ   Sequence   525 AA;

  Query Match           17.2%;  Score 357;  DB 27;  Length 525;
  Best Local Similarity  31.8%;  Pred. No. 1.93e-24;
  Matches    54;  Conservative   51;  Mismatches  57;  Indels   8;  Gaps   8;

Db      32 emcndlpynltsfpnlvdeeswkdasesiltykpllsvvcseqlkfflcsvyfpmcnekl 91
           ::|: : |    :|||:  |: |:  |   :  ||:   |  : | ||||:: |:| : |
Qy      28 QLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDL 87

Db      92 anpigpcrplclsvqekclpvlesfgfkwpdvircdkfplennrekmcmk-gpneqg-ai 149
            : | ||::||: | ::| ||: :||| |||:: ||:|| ::| : :|:  :::::  :
Qy      88 DETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP-QDN-D-LCIPLASSDHLLPA 144

Db     150 qderakf-aa-keseddgndrvediqrevdrlngkcpqdevflnrsskcv 197
           :| :|   | |: :|| || :| : ::   |: |   :  ::|| :| :
Qy     145 TEEAPKVCEACKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDTKII 193


RESULT    6
ID   W31267 standard; Protein; 694 AA.
AC   W31267;
DT   27-APR-1998  (first entry)
DE   Drosophila frizzled-2 protein (Wnt receptor).
KW   Wnt receptor; Drosophila frizzled-2 protein; Dfz2 gene;
KW   wingless receptor; Wg receptor; signal transduction; cancer;
KW   cell growth; cell proliferation.
OS   Drosophila melanogaster.
FH   Key             Location/Qualifiers
FT   Misc_difference 268
FT                   /note= "encoded by CAC"
FT   Misc_difference 269
FT                   /note= "encoded by TGG"
FT   Misc_difference 348
FT                   /note= "encoded by TA (apparent 1 nucleotide
FT                    deletion of codon)"
FT   Misc_difference 488
FT                   /note= "encoded by TTA"
FT   Misc_difference 632
FT                   /note= "encoded by CTG"
FT   Misc_difference 633
FT                   /note= "encoded by GCG"
FT   Misc_difference 671
FT                   /note= "encoded by CG (apparent 1 nucleotide
FT                    deletion of codon"
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89885.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Claim 2; Page 23-25; 61pp; English.
CC   This protein comprises the Drosophila frizzled-2 protein encoded by
CC   the Dfz2 gene (see T89885).  It is a receptor for wingless (Wg),
CC   acting as a signal transducing molecule, and is an example of a Wnt
CC   receptor (WntR).  Other novel frizzled family members have been
CC   identified in human, mouse and Caenorhabditis elegans (see
CC   W31268-74) and are considered also to be Wnt receptors.  Wnt
CC   receptors can be used in a novel, claimed method of screening for
CC   compounds which modulate the binding of a Wnt polypeptide (secreted
CC   proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC   Wnt is involved in (mammary) cancer and other processes involving
CC   growth, development and proliferation (both normal and abnormal).
CC   Modulators identified by the claimed method are useful for
CC   treatment of diseases related to these conditions.
SQ   Sequence   694 AA;

  Query Match           17.1%;  Score 355;  DB 27;  Length 694;
  Best Local Similarity  36.2%;  Pred. No. 3.07e-24;
  Matches    47;  Conservative   32;  Mismatches  45;  Indels   6;  Gaps   6;

Db      51 ygvpaip-kdpnlrceeitipmcrgigynmtsfpnemnhetqdeaglevhqfw-plveik 108
           :| | :: | :| :     : :|:|| |    :|| : |||  |  ||    | |||
Qy       8 FGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQ 66

Db     109 cspdlkfflcsmytpicledyhkplpvcrsvcerarsgcapimqqysfewpermacehlp 168
           | || | ||||:::|:||:|    :  |:|:| : :  |||:|  ::| ||: : |:::|
Qy      67 CHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFP 126

Db     169 lhgdpdnlcm 178
           : | ::||:
Qy     127 -Q-D-NDLCI 133


RESULT    7
ID   W31268 standard; Protein; 666 AA.
AC   W31268;
DT   27-APR-1998  (first entry)
DE   Mouse frizzled-3 protein Mfz3 (Wnt receptor).
```

```
KW   Wnt receptor; mouse frizzled-3 protein; Mfz3 gene;
KW   signal transduction; cancer; cell growth; cell proliferation.
OS   Mus musculus.
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89886.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Disclosure; Page 27-29; 61pp; English.
CC   This protein comprises the mouse frizzled-3 protein encoded by
CC   the Mfz3 gene (see T89886), and is a Wnt receptor.  Other novel
CC   frizzled family members have been identified in human, mouse and
CC   Caenorhabditis elegans (see W31268-74) and are considered also to
CC   be Wnt receptors.  Wnt receptors can be used in a novel, claimed
CC   method of screening for compounds which modulate the binding of a
CC   Wnt polypeptide (secreted proteins involved in cell-to-cell
CC   signalling) to a Wnt receptor.  Wnt is involved in (mammary) cancer
CC   and other processes involving growth, development and proliferation
CC   (both normal and abnormal).  Modulators identified by the claimed
CC   method are useful for treatment of diseases related to these
CC   conditions.
SQ   Sequence   666 AA;

  Query Match           15.3%;  Score 317;  DB 27;  Length 666;
  Best Local Similarity  33.3%;  Pred. No. 2.03e-20;
  Matches   44;  Conservative   29;  Mismatches  54;  Indels   5;  Gaps   5;

Db      27 scepi-t-lrmcqdlpynttfmpnllnhydqqtaalame-pfhpmvnldcsrdfrpflca 83
           :| || : |::|: : |    :|||| | :     |    :  |:|  :|  | : |||:
Qy      19 NCKPIPANLQLCHGIEYQNMRLPNLLGH-ETMKEVLEQAGAWIPLVMKQCHPDTKKFLCS 77

Db      84 lyapicme-ygrvtlpcrrlcqraysecsklmemfgvpwpedmecsrfpdcdepyprlvd 142
           |:||:|::       ||: || :    |: :|  || |||: :|| |||: ::    |
Qy      78 LFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLAS 137

Db     143 lnlvgdptegap 154
            : :  :|| ||
Qy     138 SDHLLPATEEAP 149


RESULT    8
ID   W31272 standard; Protein; 709 AA.
AC   W31272;
DT   27-APR-1998  (first entry)
DE   Mouse frizzled-6 protein Mfz6 (Wnt receptor).
KW   Wnt receptor; mouse frizzled-6 protein; Mfz6 gene;
KW   signal transduction; cancer; cell growth; cell proliferation.
OS   Mus musculus.
PN   WO9739357-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U06049.
PR   12-APR-1996; US-015307.
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI   Nusse R, Samos CH, Wangy;
DR   WPI; 97-526631/48.
DR   N-PSDB; T89890.
PT   Identification of Wnt receptor binding modulators - useful for
PT   treatment of cancer and growth, development or proliferation related
PT   disorders
PS   Disclosure; Page 41-42; 61pp; English.
CC   This protein comprises the mouse putative transmembrane receptor,
CC   frizzled-6 (Mfz6), encoded by the Mfz6 gene (see T89890).  It is
CC   an example of a Wnt receptor.  Other novel frizzled family members
CC   have been identified in human, mouse and Caenorhabditis elegans
CC   (see W31268-74) and are considered also to be Wnt receptors.  Wnt
CC   receptors can be used in a novel, claimed method of screening for
CC   compounds which modulate the binding of a Wnt polypeptide (secreted
CC   proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC   Wnt is involved in (mammary) cancer and other processes involving
CC   growth, development and proliferation (both normal and abnormal).
CC   Modulators identified by the claimed method are useful for
CC   treatment of diseases related to these conditions.
SQ   Sequence   709 AA;

  Query Match           13.7%;  Score 284;  DB 27;  Length 709;
  Best Local Similarity  33.6%;  Pred. No. 3.89e-17;
  Matches   41;  Conservative   26;  Mismatches  49;  Indels   6;  Gaps   6;

Db      24 cepitvp-r-cmkmty-nmtffpnlmghydqgiaavemghflhlanlecspniemflcqa 80
           | ||    : |  : | || : |||:||       : |  : |   :| |:  :|||
Qy      20 CKPIPANLQLCHGIEYQNMRL-PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSL 78

Db      81 fiptcteqih-vvlpcrklcekivsdckklmdtfgirwpeelecnrlphcddt-vpvtsh 138
           | | | :::   : ||: ||  :   |  :| :||: ||: |||:|:|: :|  :|::|
Qy      79 FAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASS 138

Db     139 ph 140
             |
Qy     139 DH 140


RESULT    9
ID   R74187 standard; Protein; 581 AA.
AC   R74187;
DT   24-DEC-1995  (first entry)
DE   Chick p75.
KW   Neural axon out-growth modulator; epidermal growth factor; EGF;
KW   netrin-2; p75; neurodegenerative disease; transgenic animal;
KW   gene therapy.
OS   Gallus sp.
FH   Key             Location/Qualifiers
FT   peptide         1..15
FT                   /label= Sig_peptide
FT   domain          16..260
FT                   /label= Domain-VI
FT                   /note= "domain VI is probably involved in protein-
FT                   protein complex formation"
FT   domain          261..429
FT                   /label= Domain-V
FT                   /note= "domain V includes EGF-like repeats"
FT   region          430..581
FT                   /note= "C-terminal region"
PN   WO9513367-A1.
PD   18-MAY-1995.
PF   08-NOV-1994; U12913.
PR   12-NOV-1993; US-152019.
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
PI   Dodd J,  Jessell T,  Kennedy T,  Placzek M,  Serafini T;
PI   Tessier-Lavigne M;
DR   WPI; 95-194086/25.
DR   N-PSDB; Q92367.
PT   Neural axon out-growth modulators derived from EGF-like repeats of
PT   netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
PT   increasing spinal axon out-growth or directing axon orientation
PS   Claim 1; Page 48-50; 58pp; English.
CC   An E10 chick brain cDNA library was screened with probes based on
CC   netrin-1 (p78) or netrin-2 (p75) sequences to isolate chick p78
CC   and p75 partial cDNA clones.  Full-length clones (given in
CC   Q92366-67, respectively) were subsequently obtd. by 3'RACE.  cDNA
CC   is expressed e.g. in COS or insect cells for recombinant p78 and
CC   p75 prodn., used to breed transgenic animals, or for gene therapy.
SQ   Sequence   581 AA;

  Query Match           7.1%;  Score 148;  DB 14;  Length 581;
  Best Local Similarity  27.0%;  Pred. No. 3.20e-04;
```

```
  Matches    30;  Conservative   28;  Mismatches  46;  Indels   7;  Gaps   6;

Db     431 ipainptslvtste-apadcdsyckpakgnykinmkkyckkdyvvqvnilemetvanwak 489
           ||  ::  |: :|| ||  |:: ||  : : :  |   ||:|: : ::: |:  :   :|
Qy     133 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKVKEITYINRDTK 191

Db     490 ftinils--vykcrdervkr-gdnflwihlkdlsckcpki-qiskkylvmg 536
           : ::  |  :||      :    : ||:    | | |  : :|:  |||||
Qy     192 IILETKSKTIYKLNGVSERDLKKSVLWLK-DSLQCTCEEMNDINAPYLVMG 241


RESULT   10
ID   R74186 standard; Protein; 605 AA.
AC   R74186;
DT   24-DEC-1995  (first entry)
DE   Chick p78.
KW   Neural axon out-growth modulator; epidermal growth factor; EGF;
KW   netrin-1; p78; neurodegenerative disease; transgenic animal;
KW   gene therapy.
OS   Gallus sp.
FH   Key             Location/Qualifiers
FT   peptide         1..25
FT                   /label= Sig_peptide
FT   misc_difference 29
FT                   /note= "unspecified amino acid"
FT   misc_difference 38
FT                   /note= "unspecified amino acid"
FT   misc_difference 50
FT                   /note= "unspecified amino acid"
FT   domain          26..284
FT                   /label= Domain-VI
FT                   /note= "domain VI is probably involved in protein-
FT                   protein complex formation"
FT   domain          285..453
FT                   /label= Domain-V
FT                   /note= "domain V includes EGF-like repeats"
FT   region          454..605
FT                   /note= "C-terminal region"
PN   WO9513367-A1.
PD   18-MAY-1995.
PF   08-NOV-1994; U12913.
PR   12-NOV-1993; US-152019.
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
PI   Dodd J,  Jessell T,  Kennedy T,  Placzek M,  Serafini T;
PI   Tessier-Lavigne M;
DR   WPI; 95-194086/25.
DR   N-PSDB; Q92366.
PT   Neural axon out-growth modulators derived from EGF-like repeats of
PT   netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
PT   increasing spinal axon out-growth or directing axon orientation
PS   Claim 1; Page 44-46; 58pp; English.
CC   An E10 chick brain cDNA library was screened with probes based on
CC   netrin-1 (p78) or netrin-2 (p75) sequences to isolate chick p78
CC   and p75 partial cDNA clones.  Full-length clones (given in
CC   Q92366-67, respectively) were subsequently obtd. by 3'RACE.  cDNA
CC   is expressed e.g. in COS or insect cells for recombinant p78 and
CC   p75 prodn., used to breed transgenic animals, or for gene therapy.
SQ   Sequence   605 AA;

  Query Match             6.0%;  Score 124;  DB 14;  Length 605;
  Best Local Similarity  26.1%;  Pred. No. 3.95e-02;
  Matches    31;  Conservative   34;  Mismatches  41;  Indels  13;  Gaps  11;

Db     455 ipaappptaassteepad-cdsyckaskgklkinmkkyckkdyavqihilkaeknadw-w 512
           || |::    ::|||::  |:: ||  :   :  |   ||:|:|: | : |     ::
Qy     133 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDT 190

Db     513 kftvniis--vyk-qg-snr-lrrgdqtlwvhakdiackcpkvkpmkk-ylllgsteds 565
           |: ::  |  :||  | |:| |:::   ||:    : | |  :: ::  ||::|  : :
Qy     191 KIILETKSKTIYKLNGVSERDLKKS--VLWLKD-SLQCTCEEMNDINAPYLVMGQKQGG 246
```

```
RESULT   11
ID   W35946 standard; Protein; 604 AA.
AC   W35946;
DT   11-MAY-1998  (first entry)
DE   Human netrin-1.
KW   Netrin-1; neuron; growth; differentiation; morphology;
KW   neural disease; diagnosis; therapy; drug screening; human.
OS   Homo sapiens.
PN   WO9740064-A1.
PD   30-OCT-1997.
PF   19-APR-1997; U06452.
PR   19-APR-1996; US-635137.
PA   (EXEL-) EXELIXIS PHARM INC.
PA   (REGC ) UNIV CALIFORNIA.
PI   Kennedy T, Leonardo D, Serafini T, Shyian A, Swimmer C,
PI   Tessier-Lavigne M, Zhang Y;
DR   WPI; 97-535773/49.
DR   N-PSDB; T97129.
PT   Human netrin-1 protein and related nucleic acids - useful in
PT   modulating neuron growth and screening for compounds for diagnosis
PT   or treatment of diseases associated with undesirable growth
PS   Claim 1; Page 13-15; 22pp; English.
CC   This protein comprises human netrin-1 (see W35946), a protein
CC   which is involved in neural axon guidance, and which is especially
CC   useful in modulating neural axon outgrowth.  Its amino acid
CC   sequence was deduced from a cDNA clone (see T97129) isolated from a
CC   human foetal brain cDNA library.  Neuron growth, differentiation or
CC   morphology can be altered by contact with netrin (claimed).
CC   Isolated netrin-1 can also be used to screen chemical libraries
CC   for candidate drugs suitable for diagnosis or treatment of diseases
CC   associated with undesirable neural cell growth, by comparing
CC   binding to a netrin binding target with and without the presence of
CC   a prospective agent.  Agents that modulate the interaction may be
CC   useful as pharmaceutical lead compounds (claimed).
SQ   Sequence   604 AA;

  Query Match             5.8%;  Score 121;  DB 28;  Length 604;
  Best Local Similarity  24.8%;  Pred. No. 7.10e-02;
  Matches    29;  Conservative   30;  Mismatches  49;  Indels   9;  Gaps   8;

Db     454 ipvappttaassvee-pedcdsyckaskgklkinmkkyckkdyavqihilkadkagdw-w 511
           ||:|::    :: || |  |:: ||  :   :  |   ||:|:|: | : |     ::
Qy     133 IPLASSDHLLPATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDT 190

Db     512 kftvniis--vykqgtsrirrgdqs-lwirsrdiackcpkikplkk-ylllgnaeds 564
           |: ::  |  :||      |    | ||::   : | |  :: ::  ||::|  : :
Qy     191 KIILETKSKTIYKLNGVSERDLKKSVLWLKD-SLQCTCEEMNDINAPYLVMGQKQGG 246


RESULT   12
ID   R74188 standard; Protein; 529 AA.
AC   R74188;
DT   24-DEC-1995  (first entry)
DE   Mouse p78.
KW   Neural axon out-growth modulator; epidermal growth factor; EGF;
KW   netrin-1; p78; neurodegenerative disease; transgenic animal;
KW   gene therapy.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   misc_difference 296..297
FT                   /note= "unidentified amino acids"
FT   misc_difference 300..301
FT                   /note= "unidentified amino acids"
FT   misc_difference 307..308
FT                   /note= "unidentified amino acids"
FT   misc_difference 311
FT                   /note= "unidentified amino acid"
FT   misc_difference 389
FT                   /note= "unidentified amino acid"
FT   misc_difference 393..394
FT                   /note= "unidentified amino acids"
FT   misc_difference 487
FT                   /note= "unidentified amino acid"
```

```
PN   WO9513367-A1.
PD   18-MAY-1995.
PF   08-NOV-1994; U12913.
PR   12-NOV-1993; US-152019.
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
PI   Dodd J,  Jessell T,  Kennedy T,  Placzek M,  Serafini T;
PI   Tessier-Lavigne M;
DR   WPI; 95-194086/25.
DR   N-PSDB; Q92368.
PT   Neural axon out-growth modulators derived from EGF-like repeats of
PT   netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
PT   increasing spinal axon out-growth or directing axon orientation
PS   Claim 1; Page 51-52; 58pp; English.
CC   Chick p75 and p78 sequences (given in R74186-87, respectively)
CC   were used to identify conserved amino acid regions, which were
CC   then used to design degenerate primers for the amplification
CC   of a fragment of the mouse p78 cDNA (Q92368).  This cDNA can
CC   be expressed in host cells for recombinant p78 prodn, or used
CC   to breed transgenic animals, or for gene therapy.
SQ   Sequence   529 AA;

  Query Match             5.1%;  Score 105;  DB 14;  Length 529;
  Best Local Similarity  23.5%;  Pred. No. 1.50e+00;
  Matches    24;  Conservative   25;  Mismatches  45;  Indels   8;  Gaps   7;

Db    388 sxveexxedcdsyckaskgklkmnmkkycrkdyavqihilkadkagdw-wkftvniis-- 444
          :  ||    |:: ||  :   :  |   |::|:|: | : |     ::  |: ::  |
Qy    143 PATEEAPKVCEA-CKNKNDDDNDIMETLCKNDFALKIKV-KEITYINRDTKIILETKSKT 200

Db    445 vykqgtsrirrgdqs-lwirsrdiackcpkikplkk-ylllg 484
          :||      |    | ||::   : | |  :: ::  ||::|
Qy    201 IYKLNGVSERDLKKSVLWLKD-SLQCTCEEMNDINAPYLVMG 241


RESULT   13
ID   W26642 standard; Protein; 732 AA.
AC   W26642;
DT   11-FEB-1998  (first entry)
DE   Human RECK cancer-inhibiting protein.
KW   RECK; reversion-inducing cysteine rich protein with Kazal motif;
KW   human; cancer; gene therapy; diagnosis.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          396..415
FT                   /label= Kazal_domain
PN   WO9724439-A1.
PD   10-JUL-1997.
PF   24-DEC-1996; U20812.
PR   27-DEC-1995; JP-340469.
PA   (AMGE-) AMGEN INC.
PA   (KITA/) KITAYAMA H.
PA   (NODA/) NODA M.
PA   (TAKA/) TAKAHASHI C.
PA   (SANY ) SANKYO CO LTD.
PI   Kitayama H,  Noda M,  Takahashi C;
DR   WPI; 97-363675/33.
DR   N-PSDB; T90508.
PT   RECK gene and corresponding protein sequences - enables reversion of
PT   cancer cells
PS   Claim 7; Page 42-44; 53pp; English.
CC   This protein sequence comprises the human RECK protein (reversion-
CC   inducing cysteine rich protein with Kazal motif), which is capable
CC   of transforming malignant cancer cells with an activated ras gene
CC   into normal cells (reversion activity).  Its amino acid sequence
CC   was deduced from a cDNA clone (see T90508) obtained from human
CC   fibroblast MRC-5 (ATCC CCL 171) cells.  A claimed method for
CC   treatment of cancer comprises contacting the cancer cells with a
CC   RECK polypeptide.  RECK may also be expressed using gene therapy
CC   methods for in vivo treatment of cancer.
SQ   Sequence   732 AA;

  Query Match             4.8%;  Score 100;  DB 25;  Length 732;
  Best Local Similarity  30.5%;  Pred. No. 3.78e+00;
  Matches    25;  Conservative   15;  Mismatches  36;  Indels   6;  Gaps   5;

Db    158 ipvldikkcqpemwkaiacslqikpchsksrgsii-cksdcveilkkcgd-qnkfpedht 215
          || | :| |:|:  | : |||    |      :|  | | ||::  :|:   : |
Qy     60 IP-LVMKQCHPDT-KKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMSAFGFPWP 117

Db    216 aesicellsptddlkncipldt 237
             |: ::  :||  |||| :
Qy    118 DMLECDRFPQDNDL--CIPLAS 137


RESULT   14
ID   W55368 standard; Protein; 804 AA.
AC   W55368;
DT   17-JUN-1998  (first entry)
DE   H. pylori ORF 09ap11406orf2 protein.
KW   Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW   identification; binding compound; bacteria; life cycle; activator;
KW   inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
OS   Helicobacter pylori.
PN   WO9737044-A1.
PD   09-OCT-1997.
PF   27-MAR-1997; U05223.
PR   06-DEC-1996; US-761318.
PR   29-MAR-1996; US-625811.
PR   02-APR-1996; US-758731.
PR   25-OCT-1996; US-736905.
PR   28-OCT-1996; US-738859.
PA   (ASTR ) ASTRA AB.
PI   Alm RA, Smith D;
DR   WPI; 97-503122/46.
DR   N-PSDB; V24777.
PT   Helicobacter pylori nucleic acid sequences and encoded
PT   polypeptide(s) - useful in vaccines to treat or prevent H. pylori
PT   infection and for diagnosis of H. pylori infection
PS   Claim 14; Page 587-589; 1145pp; English.
CC   This sequence is a H. pylori protein of unspecified function.
CC   The protein may be used in a vaccine to prevent or treat H. pylori
CC   infection or to identify H. pylori polypeptide binding compounds,
CC   useful as potential H. pylori life cycle activators or inhibitors. The
CC   DNA and probes derived from it may be used for the identification of
CC   H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
CC   acid sequences complementary to the DNA act as antisense sequences and
CC   can be used to prevent the translation of H. pylori mRNA. Antibodies
CC   against the protein can be used in immunoassays to evaluate the abundance
CC   and distribution of H. pylori-specific antigens. The genomic sequence of
CC   H. pylori (ATCC 55679) was determined from overlapping contigs generated
CC   by mechanically shearing the bacterial DNA. The sequences were analysed
CC   for ORF of at least 180 nucleotides, and the predicted coding regions
CC   defined by computer evaluation. To identify likely H. pylori antigens for
CC   vaccine development, the amino acid sequences predicted from various ORF
CC   were analysed for significant homology to other known or exported
CC   membrane proteins. Having identified and determined the sequences of
CC   interest, particular regions can be isolated from H. pylori by PCR
CC   amplification for recombinant polypeptide production, e.g. in E. coli
CC   hosts.
SQ   Sequence   804 AA;

  Query Match             4.8%;  Score 99;  DB 29;  Length 804;
  Best Local Similarity  31.5%;  Pred. No. 4.53e+00;
  Matches    17;  Conservative   16;  Mismatches  19;  Indels   2;  Gaps   2;

Db     78 vidsipkegftplenaffnitysmrqqqfilhikpkvmrrltlfsfdrdykkai 131
          ::::: |: |: |     :|||  |:  :||  |:| : :|   | :|| ||::
Qy    165 IMETLCKNDFA-LKIKVKEITYINRDTKIILETKSKTIYKLNGVS-ERDLKKSV 216


RESULT   15
ID   R44929 standard; Protein; 15281 AA.
AC   R44929;
DT   08-JUL-1994  (first entry)
DE   T. niveum Cyclosporin synthetase.
```

KW Enzyme; cyclosporin; synthetase-like activity; Tolypocladium niveum;
KW T. inflatum GAMS; biosynthesis; vector; cyclosporin synthetase.
OS Tolypocladium niveum.
PN EP-578616-A.
PD 12-JAN-1994.
PF 05-JUL-1993; 810474.
PR 09-JUL-1992; AT-001403.
PR 08-MAR-1993; AT-000437.
PR 29-APR-1993; CH-001310.
PR 04-MAY-1993; CH-001375.
PA (SANO) SANDOZ LTD.
PA (SANO) SANDOZ PATENT GMBH.
PA (SANO) SANDOZ-ERFINDUNGEN VERW GES MBH.
PI Leitner E, Schneider E, Schoergendorfer K, Weber G;
DR WPI; 94-010432/02.
DR N-PSDB; Q54386.
PT Isolated DNA sequence - which codes for enzyme having cyclosporin
PT synthetase like activity
PS Claim 1; Page 41-84; 93pp; English.
CC This sequence represents an enzyme which has cyclosporin synthetase-
CC like activity. This sequence was isolated from Tolypocladium niveum
CC (formerly known as T. inflatum GAMS). This enzyme catalyses the
CC peptide biosynthesis of cyclosporins and structurally related
CC molecules. This sequence may be used for the production of
CC cyclosporin by transforming a vector containing this sequence in
CC to a recombinant host. This allows effective production of anti-
CC biotic cyclosporin or its derivatives.
SQ Sequence 15281 AA;

Query Match 4.7%; Score 97; DB 9; Length 15281;
Best Local Similarity 26.7%; Pred. No. 6.52e+00;
Matches 24; Conservative 23; Mismatches 35; Indels 8; Gaps 8;

```
Db   3819 lrgtkisdhiaianipnsktiverticesvydlgg-dakdsndrvswlsaa-rsnavkva 3876
          |:  |: : |:  |  ::| |:| |   ::| | |   :| :  | ||  : :     :
Qy    176 LK-IKVKE-ITYIN-RDTKIILE-TKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCEEMN 231

Db   3877 slsaidlvdiaqeagfrveiscarqwsqng 3906
           ::|  || ::|  |  : |:  ::| |:|
Qy    232 DINAPYLV-MGQKQGGELVITSVKRW-QKG 259
```

Search completed: Thu Oct 22 16:01:16 1998
Job time : 70 secs.